

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 19:19:42 ; Search time 120 Seconds
(without alignments)
1254.185 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPEDPEETQLLGPFGSA.....CVPETKGKTLQITAHPEGR 477

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues 1017041

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mic:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Result No.	Score	Query Match	Length	ID	Description
1	2457	100.0	477	Q8WUZ9	Q8wuz9 homo sapien
2	1411.5	57.4	498	Q7ZWH3	Q7zwh3 brachydanio
3	1407	57.3	482	Q8AYP6	Q8ayp6 gallus gall
4	1167	47.5	266	Q8VDJ4	Q8vdj4 mus musculus
5	1131	46.0	248	Q8WZ05	Q8wz05 homo sapien
6	1125.5	45.8	246	Q8SPS1	Q8sps1 oviss aries
7	946	38.5	497	Q8BNT2	Q8bnt2 mus musculus
8	894	36.4	445	Q8NCC2	Q8ncc2 homo sapien
9	804.5	32.7	488	Q8MKK4	Q8mkk4 drosophila
10	789.5	32.1	506	Q8GFS9	Q8gfs9 drosophila
11	788.5	32.1	489	Q9V609	Q9v609 drosophila
12	788.5	32.1	857	Q9V608	Q9v608 drosophila
13	788	32.1	539	Q9VU17	Q9vu17 drosophila
14	787	32.0	471	Q8IQH6	Q8iqh6 drosophila
15	764.5	31.1	433	Q9V610	Q9v610 drosophila
16	736	30.0	488	Q93YP9	Q93yp9 arabidopsis

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2457	100.0	477	Q8WUZ9	Q8wuz9 homo sapien
2	1411.5	57.4	498	Q7ZWH3	Q7zwh3 brachydanio
3	1407	57.3	482	Q8AYP6	Q8ayp6 gallus gall
4	1167	47.5	266	Q8VDJ4	Q8vdj4 mus musculus
5	1131	46.0	248	Q8WZ05	Q8wz05 homo sapien
6	1125.5	45.8	246	Q8SPS1	Q8sps1 oviss aries
7	946	38.5	497	Q8BNT2	Q8bnt2 mus musculus
8	894	36.4	445	Q8NCC2	Q8ncc2 homo sapien
9	804.5	32.7	488	Q8MKK4	Q8mkk4 drosophila
10	789.5	32.1	506	Q8GFS9	Q8gfs9 drosophila
11	788.5	32.1	489	Q9V609	Q9v609 drosophila
12	788.5	32.1	857	Q9V608	Q9v608 drosophila
13	788	32.1	539	Q9VU17	Q9vu17 drosophila
14	787	32.0	471	Q8IQH6	Q8iqh6 drosophila
15	764.5	31.1	433	Q9V610	Q9v610 drosophila
16	736	30.0	488	Q93YP9	Q93yp9 arabidopsis

Query Match 100.0% ; Score 2457 ; DB 4 ; Length 477 ;
Best Local Similarity 100.0% ; Pred. No. 2.4e-157 ;

Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTPEDPEETQPLGPGGAPRRRRVFLAFAAALGFLSGFALGYSSPAISLQRAAPP 60
 Db 1 MTPEDPEETQPLGPGGAPRRRRVFLAFAAALGFLSGFALGYSSPAISLQRAAPP 60

QY 61 APRLDDAAASWFGAVVTLGAAAGVGLGGWLVDRAGKLSLLCSVPFVAGFAVITAAQDV 120
 Db 61 APRLDDAAASWFGAVVTLGAAAGVGLGGWLVDRAGKLSLLCSVPFVAGFAVITAAQDV 120

QY 121 WMLLGRLLTGLACGVASIVAPVYIETIAYPAVRGLLGSVQVMVVVGLLAYLAGWVLE 180
 Db 121 WMLLGRLLTGLACGVASIVAPVYIETIAYPAVRGLLGSVQVMVVVGLLAYLAGWVLE 180

QY 181 WRWLAVLGVCPVPSMLLMCFMPEPRFLTLTQHRQEAARLFLWGSQGHEDPPIGAE 240
 Db 181 WRWLAVLGVCPVPSMLLMCFMPEPRFLTLTQHRQEAARLFLWGSQGHEDPPIGAE 240

QY 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQQLSGVNAVVFYAETIFEBAKFKDSSLASVVGV 300
 Db 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQQLSGVNAVVFYAETIFEBAKFKDSSLASVVGV 300

QY 301 IOVLFTAVAAALIMDRAGRLLLVLSGVNVVSTSAFGAYFKLTQGGPGNSHVAISAPVS 360
 Db 301 IOVLFTAVAAALIMDRAGRLLLVLSGVNVVSTSAFGAYFKLTQGGPGNSHVAISAPVS 360

QY 361 AQPVDASVGLAWLAVGSMCLFIAGFVAGWGPIPLLMSEIPLHVKGVATGICVLTNWLM 420
 Db 361 AQPVDASVGLAWLAVGSMCLFIAGFVAGWGPIPLLMSEIPLHVKGVATGICVLTNWLM 420

QY 421 AFLVTKFSSLMELVRPYGAFWLASAFICFISVLTFLFCVPETKGTLEQITAHFEGR 477
 Db 421 AFLVTKFSSLMELVRPYGAFWLASAFICFISVLTFLFCVPETKGTLEQITAHFEGR 477

RESULT 2
 Q7ZWH3 PRELIMINARY; PRT; 498 AA.

ID Q7ZWH3
 AC Q7ZWH3
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Body;
 RA Strausberg R.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; BC049409; AAH49409.1; -
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0008643; P:carbohydrate transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR003663; Sugar transporter.
 DR Pfam; PF00083; sugar_tr; 1.
 DR PRINTS; PR00171; SUGRTRNSPORT.
 DR TIGRFAMS; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 498 AA; 54785 MW; 4458BF19DB84255D CRC64;

Query Match 57.4%; Score 1411.5; DB 13; Length 498;
 Best Local Similarity 58.0%; Pred. No. 4.9e-87;

Matches 282; Conservative 63; Mismatches 122; Indels 19; Gaps 5;

QY 2 TPEDPEETQPLGPGGAPRRRRVFLAFAAALGFLSGFALGYSSPAISLQRAAPP 61
 Db 17 TEEDRSEQDAYL-----DKVKNKGKFIATFAAVLGLSGFVLGYSSPAIPERRIQDLR 71

QY 62 PRLDDAAASWFGAVVTLGAAAGVGLGGWLVDRAGKLSLLCSVPFVAGFAVITAAQDV 121
 Db 72 LQLSVEEASWFGSVVTTIGAAAGLGLGGWIVRIGRKLSLMFCALPFIPTFTTIIAAQNH 131

QY 122 MLLGRLTGLACGVASIVAPVYIETIAYPAVRGLLGSVQVMVVVGLLAYLAGWVLE 181
 Db 132 MEYVGRVLTGLASGVTSLVVPIIISMAHERVRGTGLGSCVQVMVVVINGAYVTGLFLDW 191

QY 182 RMLAVLGVCPVPSMLLMCFMPEPRFLTLTQHRQEAARLFLWGSQGHEDPPIGAE 240
 Db 192 RMLAVASSIPPTMLLMCFMPEPRFLTLTQHRQEAARLFLWGSQGHEDPPIGAE 251

QY 235 PPIGAEQSFHLALLRQPGIYKPFIIIGVSLMAFQQLSGVNAVVFYAETIFEBAKFKDSSIA 294
 Db 252 AYKNEEQSFSLGDLKDPGVYKPLGIGVNMVLLQOFTGINAIFYAETIFEQAHFKSSDVA 311

QY 295 SVVGVVIOVLFTAVAAALIMDRAGRLLLVLSGVNVVSTSAFGAYFKLTQGGPGNSHVA 354
 Db 312 TVIVAATQVFTAAALIMDRAGRLLLVLSGVNVVSTSAFGAYFKLTQGGPGNSHVA 371

QY 355 ISAP-----VSAQPDVDSVGLAWLAVGSMCLFIAGFVAGWGPIPLLMSEIPLHVKGVAT 410
 Db 372 VLTDTHGLLEDP-----SADLAWLAVGSMGFFIAGFVAGWGPIPLLMSEIPLHVKGVAT 428

QY 411 GICVLTNWLMELVRPYGAFWLASAFICFISVLTFLFCVPETKGTLEQITAHFEGR 470
 Db 429 ALCVLTNWTCAFIVTKTFQNLMDALSSAGTWMFSLCASNVWFTAFVPTKGTLEQITAHFEGR 488

QY 471 TAHEPG 476
 Db 489 QAGFKG 494

RESULT 3
 Q8AYP6 PRELIMINARY; PRT; 482 AA.

ID Q8AYP6
 AC Q8AYP6;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Glucose transporter type 8.
 CN GLUT8.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Seki Y., Sato K., Abe H., Akiba Y.;
 RT "Glucose transporter 8 plays a major role in insulin-responsive
 RT Glucose transporter in place of glucose transporter 4 in chickens."
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB083371; BAC20934.1; -
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0008643; P:carbohydrate transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR003663; Sugar transporter.
 DR InterPro; IPR005829; Sug transporter.
 DR Pfam; PF00083; sugar_tr; 1.
 DR PRINTS; PR00171; SUGRTRNSPORT.
 DR TIGRFAMS; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.

SQ SEQUENCE 266 AA; 29132 MW; 1B3031928863B0FC CRC64;

Query Match 47.5%; Score 1167; DB 11; Length 266;
Best Local Similarity 83.1%; Pred. No. 6.9e-71;
Matches 222; Conservative 20; Mismatches 23; Indels 2; Gaps 2;

QY 212 QHRQREMAALREFLWSEGWEDPPIGAE-QSHFALLRPPGIYKPFIIIGVSLMAFOOLS 270
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 1 OHQOQEAMAAALREFLWSEGWEPPVGAHQGFQLALLRRPGIYKPLIIGISLMVFQOLS 60
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

QY 271 GVNAVIFYAETIFEEAKFKDSSLASVVGVIVQLFTAVAILMDRAGRLLLVLGGVVMV 330
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 61 GVNAIMFYANSIFEAKFKDSSLASVTGIIQVLF-TAVAALIMDRAGRLLLSLVGMV 120
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

QY 331 FSTSAFGAYFKLTQGGPGNSHVAISAPVSAQPVDASVGLAWLAVGSMCLFIAGFAVGWG 390
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 121 FMSAFPTYFKLTQSLEPSNSSHVGL-VPIAAEPDVQVGLAWLAVGSMCLFIAGFAVGWG 179
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

QY 391 PIPWLLMSEIFFLHVKGATGICVLTNWLMAFLVTKFEFSIMEVLRPYGFWFVASACIF 450
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 180 PIPWLLMSEIFFLHVKGATGCVLTNWFMAFLVTKFEFSVMEMLRPYGAFWLTAAFCAL 239
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

QY 451 SVLETLFCVPETKGTLEQTIAHEGR 477
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 240 SVLETLFWPETKGTLEQTIAHEGR 266
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 5

Q8WZ05 PRELIMINARY; PRT; 248 AA.

ID AC Q8WZ05; DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RN RA Zhang P.P., Zhou X.M., Jiang H.Q., Huang Y., Qin W.X., Zhao X.T.,
WA Wan D.F., Gu J.R.;
RT "Novel human cDNA clones with function of inhibiting cancer cell growth."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AF289587; AAL55771.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR005829; Sug transporter.
DR Pfam; PF00083; sugar tr; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
KW Hypothetical protein; Transmembrane.
SQ SEQUENCE 248 AA; 27017 MW; AA929A7499BEE558 CRC64;

Query Match 46.0%; Score 1131; DB 4; Length 248;
Best Local Similarity 93.0%; Pred. No. 1.7e-68;
Matches 227; Conservative 5; Mismatches 6; Indels 6; Gaps 2;

QY 164 MVVVGILLAYLAGVLEWRNLAVLCVPPSPMLLLMCMFPETPRFLTQHRRQEAVALR 223
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 1 MVVVGILLAYLAGVLEWRNLAVLCVPPSPMLLLMCMFPETPRFLTQHRRQEAVALR 60
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

QY 224 FLWSEQGWEDPPIGAESFHIALIRQGIYKPFIIIGVSLMAFOOLS 283
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 61 FLWSEQGWEDPPIGAESFHIALIRQGIYKPFIIIGVSLMAFOOLS 120
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

QY 284 EEAKFKDSSLASVVGVIVQLFTAVAILMDRAGRLLLVLGGVVMVFTSFAFGVKLT 343
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
NON TER

Db 121 EBAKFKDSSLASVVGVIQVLTAVAAALIMDRAGRRLLVLGVMVVFSTSAFGAYFKLT 180
 QY 344 QGGPGNSHVAISAPVSAQPDVDSVGLAWLAVGSMCLFTAGFVAGWGP-IPWLLMSEIFP 402
 Db 181 QGGPGNSHVAISAPVSAQPDVDSVGLAWLAVGSMCLFTAGFVAGWGP-IPWLLMSEIFP 402
 QY 403 LHVK 406
 Db 236 LHLQ 239

RESULT 6

Q8SPS1 PRELIMINARY; PRT; 246 AA.
 AC Q8SPS1;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Glucose transporter 8 (Fragment).
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 CX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Linesand S.W., Regnault T.R.H., Hay W.W. Jr.;
 RT "Characterization of Glucose Transporter 8 (GLUT8) in the Ovine Placenta".
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 DR EMBL; AF495799; AAM18513.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0008643; P:carbohydrate transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR003663; Sugar transport.
 DR InterPro; IPR005829; Sug transporter.
 DR Pfam; PF00083; sugar tr; 1.
 DR PRINTS; PR00171; SUGRTNSPORT.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR TRANSMEMBRANE.
 FT NON TER
 SQ SEQUENCE 246 AA; 26537 MW; E8F3871946CD0723 CRC64;

Query Match 45.8%; Score 1125.5; DB 6; Length 246;
 Best Local Similarity 88.2%; Pred. No. 3.9e-68;
 Matches 217; Conservative 11; Mismatches 17; Indels 1; Gaps 1;
 QY 219 MAALRFLWGSQEQWEDPPTGAE-QSHALLRQPGYKPIIGVSLMAFQOLSGVNAVVF 277
 Db 1 MAALRFLWGSQEQWEDPPTGAE-QSHALLRQPGYKPIIGVSLMAFQOLSGVNAVVF 60
 QY 278 YAEITFEAKFKDSSLASVVGVIQVLTAVAAALIMDRAGRRLLVLGVMVVFSTSAFG 337
 Db 61 YAEITFEAKFKDSSLASVVGVIQVLTAVAAALIMDRAGRRLLVLGVMVVFSTSAFG 120
 QY 338 AYFKLTQGGPGNSHVAISAPVSAQPDVDSVGLAWLAVGSMCLFIAGFVAGWGP-IPWLLM 397
 Db 121 AYFKLTQGGPGNSHVAISAPVSAQPDVDSVGLAWLAVGSMCLFIAGFVAGWGP-IPWLLM 180
 QY 398 SEIFPLHVKGATGICVLTNMLMAFLVTKFESLSMEVLVDPYGAFWLASFCIFSVLFTLF 457
 Db 181 SEIFPLHVKGATGICVLTNMLMAFLVTKFESLSMEVLVDPYGAFWLASFCIFSVLFTLF 240
 QY 458 CVPETK 463
 Db 241 CVPETK 246

RESULT 7

Q8BTN2 PRELIMINARY; PRT; 497 AA.
 AC Q8BTN2;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Similar to solute carrier family 2.
 GN F630103L12RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573 (2002).
 DR EMBL; AK089246; BAC40811.1; -;
 DR MGD; MGI:2443286; F630103L12RIK.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0008643; P:carbohydrate transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR003663; Sugar transport.
 DR InterPro; IPR005829; Sug transporter.
 DR Pfam; PF00083; sugar tr; 1.
 DR PRINTS; PR00171; SUGRTNSPORT.
 DR TIGRFAMs; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 DR NON TER
 SQ SEQUENCE 497 AA; 54445 MW; 8E14D1DCD4B244E3 CRC64;
 Query Match 38.5%; Score 946; DB 11; Length 497;
 Best Local Similarity 43.8%; Pred. No. 9.6e-56;
 Matches 214; Conservative 77; Mismatches 168; Indels 32; Gaps 9;
 QY 3 PEDPEETQPLLGPFG---GSAPGRRRVFLAAFAAALGPLSGFALGYSSPAIPSLQRAA 58
 Db 17 PEVP-----ATPGERERAGALKNRVFLATPAVLGNFSGYALVYTSFVPIELKLS 69
 QY 59 PPAFLRDDAAASWFGAVVTILGAAAGVGLGWLVDRAGRKLSLLCSVPFVAGFAVITAAQ 118
 Db 70 DPALHLDKIQASWFGSVFTLGAAGAGLSAMLLNDLLGRKLSIMFSAPSAIGVIMAGAR 129
 QY 119 DVWMLLGGRLNLGLACGVASLVAPVVISIATPAVAGLILGSCVOLWVVGILLAYLAGWV 178
 Db 130 GLWMLLGGRLNLGLACGVASLVAPVVISIATPAVAGLILGSCVOLWVVGILLAYLAGWV 189
 QY 179 LEWRMLAVLGCVPSPSLMILLMCMFETPRFLITQHRHQEAMAAALRLFWG-SEQGWEDPPI 237
 Db 190 LEWRMLAVAGEPVLIMILLMCMFETPRFLITQHRHQEAMAAALRLFWG-SEQGWEDPPI 249
 QY 238 G---AEQSHL--ALLRQPGIKYKPIIGVSLMAFQOLSGVNAVVFYAEITFEAK-FKDS 291
 Db 250 QNVNRQSRVSWAEAREPVRPVLIAVIMRFLQLTGITLPLVVLQITFDNTSVVLPFS 309
 QY 292 SLASVVGVIQVLTAVAAALIMDRAGRRLLVLGVMVVFSTSAFGAYFK-----LTQGG 346
 Db 310 QQDAIVAGVRLLSVLLIAAVTMDLGRKVLVLSVASVMEFAANITLGLYVQVPRPJT--- 366
 QY 347 PGNSSHVAISAPVSAQPDVDSVGLAWLAVGSMCLFIAGFVAGWGP-IPWLLMSEIFP 406
 Db 367 PNSTVEIVILGDT-----AFNYLTILPLLATMFLMGVGMGPITWLLMSEVLP 420

```
QY 407 GVATGCVLTNWMFLVTKERSSLMELVRPYGAFWLASAPCLFSLFVFCVPETKGT 466
Db 421 GVASGLCVLSWLTAFVLTNYFLAVNAFGLQVPFFFAICLLSLFTGCCVPETGRS 480
QY 467 LEQITAHREGR 477
Db 481 LEQIEAFFTR 491

RESULT 8
Q8NCC2 PRELIMINARY; PRT; 445 AA.
AC Q8NCC2;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ90355.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RA "NEDO human cDNA sequencing project.";
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RL -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
EMBL; AK074836; BAC1235.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; F:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR003663; Sugar transp.
DR InterPro; IPR005829; Sug transporter.
DR Pfam; PF00083; sugar tr-1.
DR PRINTS; PR00171; SUGTRNSPT.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR TRANSPORT 1; 1.
DR PROSITE; PS00217; SUGAR TRANSPORT 2; 1.
KW Hypothetical protein; Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 445 AA; 48040 MW; BF37DC0C313A32CE CRC64;

Query Match 36.4%; Score 894; DB 4; Length 445;
Best Local Similarity 40.1%; Pred. No. 2.7e-52;
Matches 200; Conservative 73; Mismatches 134; Indels 92; Gaps 9;

QY 10 QPLIG-----PPGSAAPRGR-----RVFLAAPAALGPLSGFALGYSSPA 50
Db 3 EPLLGAAGSPDYTPPEKPPSPGDRARVGLTQNRVFLATPAVLGNFSGFALVYSPV 62
QY 51 IPSLQRAAPPAPRLDDAAAGWFAVVTIGAAAGVGLGWLVDRAGRKLSLLCSVPFVAG 110
Db 63 IPALERSLDPDLHLTKQASWFGSVFTLGAAGGLSAMLNDLGRKLSIMFSAVPSAAG 122
QY 111 PAVTAAQDVWMLLGGRLTLTGLAGVASLVAAPVYSEIYAPAVRGLGSCVQLMVVVGIL 170
Db 123 YALMAGAGHLLWLLIGRTLTLTGAGGLTAACIPVYVYSEIAPPVGRGALGATPQLMAVFGSL 182
QY 171 LAYLAGVLEWRVLAVLGVCPSPMLLMCMFMPETPRELLTQHRQRQAAALRLWQSEQ 230
Db 183 SLYALGLLPLWRVLAVAGEAPVLIMLLLSFMPNSPRELLSRGDEERALRALWLRGIDV 242
QY 231 G--WEDPPIG---AEQSPHL--ALLRQPIKPFITIGVSLMAFQOLSGVNAVMFYAETIF 283
Db 243 DVHWEFEQIQNVRRQSRQSVSWAERAPHVCEPITVALLMRLQLLTGTLVYLQISIF 302
```

```
QY 284 BEAKP-----KOSLSASVVVGVLQVLTAAVALIMDRAGRELLLVLSGVVWFSTSAFGAY 339
Db 303 DSTAVLLPKPD---AAIVGAVRLLSVLTAULTMDLAGRKVL----- 341
QY 340 FKLTOGGPGNSHVAISAPVSAQPVDASVGLAWLAVGSMCLFIAGFAVGWGPIMPILMSE 399
Db 342 -----LFVSGYAVGWGPITWLMSE 361
QY 400 IPFLHVKGVATGICVLTNWMFLVTKERSSLMELVRPYGAFWLASAPCLFSLFVFCV 459
Db 362 VLPLRARGVASGLCVLASWLTAFVLTNPLVSTFGLQVPFFFAICLLSLVFTGCCV 421
QY 460 PETKGTLEQITAHRE-GR 477
Db 422 PETKGRSLQIESFFTRGR 440

RESULT 9
Q8MKK4 PRELIMINARY; PRT; 488 AA.
AC Q8MKK4;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE AT19440P (CG8234-PA).
GN CG8234.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RT Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
STRAIN=Berkely.
MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.I., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
```

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., Smith H.O.,
RT "The genome sequence of *Drosophila melanogaster*,"
RL Science 287:2185-2195 (2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B., Schemler F.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirska R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
RA "Sequencing of *Drosophila melanogaster* genome,"
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,
RA "Annotation of *Drosophila melanogaster* genome,"
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.,
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: AY122079; AAM52591.1; -;
DR EMBL: AE003825; AAM68715.1; -;
DR FlyBase; FBgn0033644; CG8234.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR003663; Sugar transp.
DR InterPro; IPR005829; Sug transporter.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRfams; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 488 AA; 53088 MW; 6554C96EBE8A1C08B CRC64;
Query Match 32.7%; Score 804.5; DB 5; Length 488;
Best local similarity 32.4%; Pred. No. 3.1e-46;
Matches 173; Conservative 81; Mismatches 177; Indels 31; Gaps 5;
QY 28 LAFAAALGSLSGFALGYSPPALPSLQRAAPPAPRLDDAAASWFGAVNVTGAAAGVGLG 87

Db 30 LAALSUSLCSLVVGFVSAYTSPALVSMTRITTSFEVTKDAGSWVGIMPLAALAGGTG 89
QY 88 GMLVDRAGRLKSLLLCSVPFFVAGFAVITAAQDVMMLLGRLTLGLACGVASIVAPVYISE 147
Db 90 GFLIEYLGRRSTILATAPFFIVSSLLIACAVNVIMILGRLTGFVGVASLSLPVYLGE 149
QY 148 IAYPAVRGLLSCVOLMVVVGILLAYLAGWLEWRWLAVLCVPPSLMLLMCFMPEPTR 207
Db 150 TLQPEVRGTGLLPTALGNIGILVCYVAGSFMMNSMLAFLGAALPVPELILMIITPETER 209
QY 208 FLITQHRQEQEAMALRLPLWSE-----QGWEDPPICASQSHALLRQGIYKP 256
Db 210 WFNRGGEERARKALKMLRGKEADVEPELKMOSQADADQATQNTCLTLFKRNL-KP 268
QY 257 FIIGVSLMAFOQLSGVNAVIFYAETIFEEA-KFQDSSLASVWVGIVQLVFAVALIMDR 315
Db 269 LSIISGLMFFQOFGSINAVIFTYQIFKDGSTIDSNLSTIIVGVNFFATFMGILLDR 328
QY 316 AGRLLLVLSGVVWVFTSAPGAYFKULTQGGPGNSHVAISAPVSAQPVDSVGLAMLV 375
Db 329 LGRKILLYSDIAMIVTSLISLGFFCYCAHGP-DVSH-----LGMPL 370
QY 376 GSMCLTAGFAGVWGPIPLMLMSEIFPLHVKGATGICVLNMLMAFLVTKERSLMEVL 435
Db 371 TCFVIYILGSLGFGPIPLMGMGBILPAKIRGPAASVVTAFNWFCTFVVTKFQDLTVAM 430
QY 436 RPYGAFWLASAFICFISVLTFLFCVPETKGTLEQITAHFEGR 477
Db 431 GAHGAFWFGAICIVGLFFVIFVPEVRGKSLERIKMWR 472

RESULT 10
ID Q86P59 PRELIMINARY; PRT; 506 AA.
AC Q86P59;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RH04286P.
GN CG30035.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y;
RA Stapleton M., Brokstein P., Hong L., Agbanyani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT003466; AAC39469.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR003663; Sugar transp.
DR InterPro; IPR005829; Sug transporter.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRfams; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
SQ SEQUENCE 506 AA; 55235 MW; 24E8718260D4319B CRC64;

Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=107311132;
 RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Flandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.U.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Borchan M.R., Bouck Y., Brockstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Hristin D., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Markulov G., Malshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson X.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Follard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., Smith H.O.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Celnikier S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Ranzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson X.A., Nunoo J.,
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phohanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
 RT "Sequencing of *Drosophila melanogaster* genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman J.B., Carlson J.W., Celnikier S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Millburn G., Richter J.F., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,
 RT "Annotatation of *Drosophila melanogaster* genome."

[illegible]

OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	Db	15	AGSNWGGIMPLAALAGGTTGGPLTEYLGRRTILATAVPFIVSSLLITACAVNMILCGR	74
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	QY	128	LTGLACGVASLVPVYISEIAYPAVRGLIGSCVOLMVVVGILLAYLAGWLEWRMLAVL	187
OC	Ephydroidea; Drosophilidae; Drosophila.	Db	75	FLTGFCVGIASLSLPVLGETLQPEVRGTLGLPTALGNITGILVCYVAGSFMMWMLAFL	134
OC	NCBI_TaxID=7227;	QY	188	GCVPSPMLLLMCWMPETERPELLTCHRRQZMAALRFLWGE-----QGWRDPP	236
RC	SEQUENCE FROM N.A.	Db	135	GAALPVPEFLIIMIIIPETPRFWNRGQEEERAKALKWLRGEADVPELKELMQSOADAD	194
RC	STRAIN=Berkley;	QY	237	ICAEQSFHALLRQPGIYKPIIIGVSLMAFOQLSGVNAMFYAETIFEEA-KFKDSSLAS	295
RA	Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,	Db	195	RQATQNTCLLELFKRNKL-KPLSISLGLMFFQOQFSGINAVIYTVQIEFKDAGSTIDSNLST	253
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,	QY	296	VVGVQIVQVLTAVAALIMDRAGRELLLVLSGVVMVFSTSAFGAFKLTQCGPGNSSHVAI	355
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,	Db	254	IIVGVNNEFAIFMGILLIDRLGRKILLYVSDIAMIVTILSLIGGFYCKAHGP-DVSH---	309
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,	QY	356	SAPVSAQPVDSAGVLAWLAVGSMCLFIAGRAVGMGPPLWLMSSIFFLHVKGATGICVL	415
RA	Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,	Db	310	-----LGMPLTCEFYIILGSLGFGPIPLWMMGBILPAKIRGPAASVVUTA	355
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,	QY	416	TNWLMAFLVTKPEFSSLMELVPYCAPWLASAFCLFSVLFTFCVPEYKGTLLQITAHPE	475
RA	Abril J.F., Agapayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,	Db	356	ENWFCTFVYTKFTQDLTVAMGAHGAFMWLFGAICIVGLFFVFIIVFPETRGKSLIEERKM	415
RA	Baillet R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,	QY	476	GR 477	
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,	Db	416	GR 417	
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,	QY			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,	Db			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,	QY			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,	Db			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,	QY			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,	Db			
RA	Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,	QY			
RA	Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,	Db			
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,	QY			
RA	Hosdin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,	Db			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,	QY			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,	Db			
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,	QY			
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,	Db			
RA	McKulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,	QY			
RA	Mount S.M., Mov M., Murby B., Murphy L., Muzny D.M., Nelson D.L.,	Db			
RA	Murphy B., Murby B., Murphy L., Muzny D.M., Nelson D.L.,	QY			

Search completed: September 27, 2004, 19:29:18

Job time : 123 secs

Search completed: September 27, 2004, 19:29:18
Job time : 123 secs

```
Query Match          31.1%; Score 764.5; DB 5; Length 433;  
Best Local Similarity 38.2%; Pred. No. 1.3e-43;  
Matches 161; Conservative 74; Mismatches 156; Indels 31; Gaps 5;  
  
QY      68 AASWFGAVVTGLGAAGGVLTGWLVDVRAGRKLSLLLCSPFPFVAGFAVITAAQDVVMLLCGR 127  
      ||| : : | ||| : || : : ||| : ||| : ||| : ||| : ||| : ||| : |||
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 19:29:30 ; Search time 147 Seconds
(without alignments)
1043.425 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPEDPEETQPLIGPPGSA.....CVPETKGTKLEQITAHFGR 477

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2457	100.0	477	14	US-10-168-651-1
2	2450	99.7	477	12	US-09-886-954-1
3	1491	60.7	326	15	US-10-254-237-2093
4	1135	46.2	262	12	US-10-169-395-9
5	948	38.6	507	14	US-10-157-031-359
6	948	38.6	507	16	US-10-755-889-110
7	723	29.4	501	13	US-10-051-909-38
8	716	29.1	484	12	US-10-424-599-275264
9	702	28.6	501	16	US-10-437-963-163544
10	702	28.6	523	12	US-10-425-114-63431
11	701.5	28.6	509	12	US-10-425-114-59933
12	658.5	26.8	553	13	US-10-051-909-36
13	639	26.0	481	15	US-10-310-154-725
14	615	25.0	488	12	US-09-794-822-11
15	615	25.0	488	13	US-10-094-059-4

16	615	25.0	488	14	US-10-170-528-5
17	615	25.0	488	14	US-10-162-012-46
18	615	25.0	488	14	US-10-062-960B-4
19	615	25.0	488	14	US-10-144-624-4
20	615	25.0	488	15	US-10-162-102-46
21	615	25.0	488	16	US-10-038-854-170
22	613.5	25.0	487	9	US-09-795-693-27
23	613.5	25.0	487	13	US-10-095-139-14
24	613.5	25.0	487	14	US-10-156-239-27
25	613.5	25.0	487	14	US-10-199-485-27
26	613.5	25.0	487	15	US-10-391-399-45
27	609.5	24.8	502	15	US-10-437-963-166774
28	607	24.7	461	15	US-10-369-493-23371
29	604.5	24.6	592	12	US-10-425-114-63791
30	604.5	24.6	596	16	US-10-437-963-177463
31	604.5	24.6	601	12	US-10-425-114-63432
32	603	24.5	472	13	US-10-024-623-29
33	603	24.5	472	14	US-10-154-419-79
34	603	24.5	472	14	US-10-146-733-74
35	603	24.5	472	15	US-10-369-493-23536
36	598.5	24.4	465	12	US-10-282-122A-56366
37	598	24.3	486	9	US-09-860-232A-7
38	593.5	24.2	525	12	US-10-425-114-63426
39	591.5	24.1	464	12	US-10-282-122A-42918
40	591.5	24.1	464	13	US-10-024-623-28
41	591.5	24.1	464	14	US-10-154-419-78
42	591.5	24.1	464	14	US-10-146-733-73
43	591.5	24.1	464	15	US-10-369-493-23552
44	591.5	24.1	523	16	US-10-437-963-124519
45	588.5	24.0	535	9	US-09-795-693-20

ALIGNMENTS

RESULT 1

US-10-168-651-1
; Sequence 1, Application US/10168651
; Publication No. US20030171275A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BURFORD, Neil
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: YANG, Junning
; APPLICANT: REDDY, Roopa
; APPLICANT: LAL, Preeti
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Valda
; APPLICANT: YUE, Henry
; APPLICANT: NGUYEN, Dannel B.
; APPLICANT: YAO, Monique G.
; APPLICANT: GANDHI, Ameeni R.
; APPLICANT: TANG, Y. Tom
; APPLICANT: KHAN, Farrah A.
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
; FILE REFERENCE: PI-0005 PCT
; CURRENT APPLICATION NUMBER: US/10/168,651
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: 60/172,000; 60/176,083; 60/177,332; 60/178,572; 60/179,758
60/181,625
; PRIOR FILING DATE: 1999-12-23; 2000-01-14; 2000-01-21; 2000-01-28; 2000-02-02;
2000-02-10
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030171275A1 1416107CD1

Sequence 5, Appli
Sequence 46, Appl
Sequence 4, Appli
Sequence 4, Appli
Sequence 46, Appl
Sequence 170, App
Sequence 27, Appl
Sequence 14, Appl
Sequence 27, Appl
Sequence 27, Appl
Sequence 45, Appl
Sequence 166774,
Sequence 23371, A
Sequence 63791, A
Sequence 177463,
Sequence 63432, A
Sequence 29, Appl
Sequence 74, Appl
Sequence 23536, A
Sequence 56366, A
Sequence 7, Appli
Sequence 63426, A
Sequence 42918, A
Sequence 28, Appl
Sequence 78, Appl
Sequence 73, Appl
Sequence 23552, A
Sequence 124519,
Sequence 20, Appl

60.7%; Score 1491; DB 15; Length 326;

	Best Local Similarity	96.3%;	Pred. No. 3.2e-123;				
	Matches	289;	Conservative	1;	Mismatches	10;	
				Indels	0;	Gaps	0;
Qy	164	MVVVGILLAYLAWVLEWRRLAVLGCVPSPSMLLLMCFMPTPRFLLTQHRQRQAMALR	223				
Db	1	MVVVGILLAYLAWVLEWRRLAVLGCVPSPSMLLLMCFMPTPRFLLTQHRQRQAMALR	60				
Qy	224	FLMGSEQGWEDPPIGAEQSFHALLRLOFGIYKPTIIGVSLMAFQOLSGVNAVMFYAETIF	283				
Db	61	FLMGSEQGWEDPPIGXEQSFHALLRLOFGIYKPTIIGVSLMAFQOLSGVNAVMFYAETIF	120				
Qy	284	EEAKFKDSSLASVVVGVIQVLFITVAALIMDRAGRLLLVLSGVVMVFSTSAFGAYFKLT	343				
Db	121	EEAKFKDSSLASVVVGVIQVLFITVAALIMDRAGRLLLVLSGVVMVFSTSAFGAYFKLT	180				
Qy	344	QGGPGNSSHVAISAPVSAQPPVDASVGLAWLAVGSMCLFIAGFVAGWGIPWLLMSEIEPL	403				
Db	181	QGGPGNSSHVAISAPVSAQPPVDASVGLAWLAVGSMCLFIAGFVAGWGIPWLLMSEIEPL	240				
Qy	404	HVKGVATGICVLTNMLMAFLVTKFSSLMVELRYPYGAFWLASAFCIFSVLFTFLFCVPETK	463				
Db	241	HVKGVATGICVLTNMLMAFLVTKFSSLMVELRYPYGAFWLASAFCIFSVLFTFLFCVPETK	300				

```

RESULT 4
US-10-169-395-9
; Sequence 9, Application US/10169395
; Publication No. US20040034192A1
; GENERAL INFORMATION:
; APPLICANT: KATO, Seishi
; APPLICANT: KIMURA, Tomoko
; TITLE OF INVENTION: HUMAN PROTEINS HAVING HYDROPHOBIC DOMAINS AND DNAs ENCODING
; TITLE OF INVENTION: THESE PROTEINS
; FILE REFERENCE: 01997.015100.US
; CURRENT APPLICATION NUMBER: US/10/169,395
; CURRENT FILING DATE: 2002-11-29
; PRIOR APPLICATION NUMBER: JP 2000-585
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: JP 2000-588
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: JP 2000-2299
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-26862
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: JP 2000-58367
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/JP00/09359
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 150
; SEQ ID NO 9
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-169-395-9

```

```

Db      181  WRWLAVLGVCPSPMLLLMCFMPTETPRFLITQHRREQAAPGLVRCGHGVQHECLRLLLQA  240

QY      229  EQCW 232
        :  ||
Db      241  DPGW 244

RESULT 5
US-10-157-031-359
; Sequence 359, Application US/10157031
; Publication No. US20030108890A1
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashev, A. V.
; APPLICANT: Krutovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US/10/157,031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 359
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-157-031-359

```

RESULT 6
US-10-755-889-110
: Sequence 110, Application US/10755889

Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 110
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-755-889-110

Query Match 38.6%; Score 948; DB 16; Length 507;
Best Local Similarity 42.7%; Pred. No. 5,5e-75;
Matches 216; Conservative 80; Mismatches 166; Indels 44; Gaps 11;

```
QY 10 QPLLG-----PPGSAAPGR-----RVFLAFAAALGPLSFGFALGYSSPA 50
D 3 EPLLAGEGPDYDTPEKPPSPGDRARVGTTLQNKRVFLATFAAVLGNFSGYALVYTSVP 62
QY 51 IPSLQRAAPPRLDDAASFGAVVTILGAAGGVGLGWLVDRAGRKLSLLCSVPFVAG 110
D 63 IPALERSLDPLHLTKSOASFGSVFTLGAAGAGLSAAMLNDLGRKLSIMFSAVPSAAG 122
QY 111 FAVITAAQDVWMLLGGRLTLGLACGVASLVAPVYISETIAYPAVRGLLGSCVQLMNVVGL 170
D 123 YALWAGAHGLWMLLGLTLTGAGLTAACTPVTVSEIAPCGVAGALGATPQLMAVFGSL 182
QY 171 LAYLAGVWLEWRLAVLCVPPSLMLLMCFMPETPRFLTLQHRROEAMALRFLMGSEQ 230
D 183 SLVAGLGLLPLWRLAVAGEADVLMLLSMPNSPRFLLSRGDRDEALRALAWLRGTDV 242
QY 231 G--WEDPPIG--AEQFHL--ALLRPGIYKPIIGVSLMAFOOLSGVNAVMEYATIF 283
D 243 DVHWEFQIQNVRRQSRVSWABARAPHVCRPTITVALLMRLQLQLTGITPILVYLSIF 302
QY 284 BEAKF---KDSLSASVWVGVIQVLTAVAAALMDRAGRLILVLSGVWVVFSTFARGAY 339
D 303 DSTAVLLPPKDD---AAIVGAVRLLSVLIAALTMDLAGRKVLLFVSAIMFANLTLGLY 359
QY 340 FKLTQGGPGNSHVAISAPVS-----AQVDSVG--LAWLAVGSMCLFIAGFAVGWGP 392
D 360 IHF---GPRPLSPNSTAGLESSEMGDLAQPLAAPAGVLTLPVLLATMLFTMGVAVGWGP 416
QY 393 PWLLMSBIFPLHVKGATGICVLTNWLMFAFLVTKFSSLMELVLPYCAFMLASAFCLPSV 452
D 417 TWLLMSVLEPRARGVAGSLCVLASMLTAFVLTKSFLPVVSTGLQVPPFFFAAICLVSL 476
QY 453 LFTLFCVPTKGTLEQITAHFE-GR 477
D 477 VFTGCCVPTKGRSLEQIESFFMRGR 502
```

RESULT 7

US-10-051-909-38
; Sequence 38, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Heientjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP

; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 38
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-051-909-38

Query Match 29.4%; Score 723; DB 13; Length 501;
Best Local Similarity 35.6%; Pred. No. 4.2e-55;
Matches 173; Conservative 89; Mismatches 172; Indels 52; Gaps 8;

```
QY 17 GGSAPRRRRVFLAFA-----AALGLSFGFALGYSSPAISLQRAAPPAP 62
D 39 GGGGMSRLGSSAYSRLDSSVAVLCTLIIVALGPIQFGFTCGFSPT----QDAISDL 94
QY 63 RLDDAASWFGAVVTILGAAGGVGLGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDVW 122
D 95 GLTLESEFSLFGSLSNVGAMVGAIAAGQIAEYIGRKSLMIAAIPNIIGWLAIKSDSF 154
QY 123 LLGRLTLGLACGVASLVAPVYISETIAYPAVRGLLGSCVQLMNVVGLLILAYLAGVLEWR 182
D 155 LFMRLLEGFGVGVISYVVPVYIETIAPQTMRGALGSVQLSVTIGILLAYLLGMFVWR 214
QY 183 WLAVLGCVPSPSLMLLMCFMPETPRFLTLQHRROEAMALRFLMGSEQMEDPPIGAE-- 240
D 215 ILSVLGILPCSILLPGIFFIPESPRWLAKMKMEDFESSLOVLRGFE-----TDIAEVN 269
QY 241 -----QSFHIALLRQPGIYKPIIGVSLMAFOOLSGVNAVMEYATIFEEAKF 288
D 270 EIKESVSSRRRTTIREADIKQKRYSVPLVMVGIGLLVLQQLSGVNGILFVAASIFKAAGL 329
QY 289 KDSLSASVWVGVIQVLTAVAAALMDRAGRLILVLSGVWVVFSTFARGAYFKLTQGGPG 348
D 330 TSNLATFGVGVQVVAIGVTTWLTDKAGRRLLIITTTGTTITLVVVSF-FVKONIT 388
QY 349 NSSHVAISAPVSAPQVDSVGLAWLAVGSMCLFIAGFAVGWGPPIWLLMSBIFPLHVKG 408
D 389 NGSHL-----YSV-MSMLSLVGLVAFVISELSGLGAIPWIMSEILPVNIKSL 435
QY 409 ATGICVLTNWLMFAFLVTKFSSLMELVLPYCAFMLASAFCLFVLTFCVPTKGTLE 468
D 436 AGSVATLANWLTAWLTMT-ASLMLSNWNGGTFAIYAAVCAAGTLVFCVLPVPTKGTLE 494
QY 469 QITAHF 474
D 495 EIAFSF 500
```

RESULT 8

US-10-424-599-275264
; Sequence 275264, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 275264
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_90584C.1.pap
US-10-424-599-275264

Query Match 29.1%; Score 716; DB 12; Length 484;
Best Local Similarity 35.6%; Pred. No. 1.7e-54;
Matches 164; Conservative 89; Mismatches 174; Indels 34; Gaps 7;

QY 26 VFLAFAAALGPIISFGFALGYSSPAIPSLQRAAPPAPRLDDAAASWFGAVVTILGAAAGGV 85
DB 45 VLFCLVIALGPIQFGFTGCGYSSPTQGAIRDL-----NLSEFSFSGSLSNVGMVGA 100
QY 86 LGGLVDRAGKLSLLCSVPFVAGFAVITAADQVWMLLGGRLITGLACGVASIVAPVYI 145
DB 101 ASQIAEYIGRKSLMIAAIPNIIGWLAIISFAKDSFLYMGRLLEGFGVGIISVVPVYI 160
QY 146 SETIAPVAVRGLGSCVQLMVVGGILLAYLAGWLEWRWLAVILGCVPSLSMLLMCFMPE 205
DB 161 AETAPQNLRGGLGSLVNLQSVTIGIMLAYLLGLFVNMVRLAILGILPCTVLIPGLFFIPES 220
QY 206 PRELLTQHRROEAMAAALRFLWGSEQGWEDPPI-----GAEQSFLHALLRQPGI 253
DB 221 PRLAKNMMDDEFTSLQVLRGFE-----TDISEVNDIKRAVASANKRTTIRFOELNOKK 293
QY 254 YKPFPIIGVSLMAFOQLSGVNAVMEYAEITFEAKFKDSSLASVVGVIQVLTAVAAALIM 313
DB 278 WFLMVIGILLVLOQLSGINGILPYSTIFANAGISSSEAAATVGLGAVQVIATGISTWLV 337
QY 314 DRAGRRLLLVSGVVMVFSTAFGAYFKLTQGGPGNSHVAISAPVSAQPVDSVGLAWL 373
DB 338 DKSGRRLLLISSSMVMSVLLIVSIAFYLL-EGVVSDESHL-----FSILGIVSVGLVAM 391
QY 374 AVGSMCLFIAGFVAGWGPPIKMLMSEIFPLHVKVATGICVLTNMLMAFLVTKFSSLM 433
DB 392 VI-----GFSGLGPIPLMIMSEILPVNIKLAGSIATMGWNLISWGITMTANLLN 443
QY 434 VLPRYGAFWLASAFICFISVLTFLFCVPETKGTLEQITAHF 474
DB 444 -WSGGFTTIVTVAATIAFIAMVVPETKGTLEBIIQFSF 483

RESULT 9
US-10-437-963-163544
; Sequence 163544, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 163544
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_62530C.1.pap
US-10-437-963-163544

Query Match 28.6%; Score 702; DB 16; Length 501;
Best Local Similarity 35.9%; Pred. No. 3e-53;
Matches 166; Conservative 80; Mismatches 178; Indels 38; Gaps 6;

QY 27 FLAFAAALGPIISFGFALGYSSPAIPSLQRAAPPAPRLDDAAASWFGAVVTILGAAAGGV 86
DB 85 FLCTLIVAGPIQFGFTGCGYSSPTQDAIRDL-----LTLSEFSVFGSLSNVGMVGA 140
QY 87 GGLVDRAGKLSLLCSVPFVAGFAVITAADQVWMLLGGRLITGLACGVASIVAPVYI 146
DB 141 SGQMAEYIGRKSLMIAAIPNIIGWLAIISFAKDSFLYMGRLLEGFGVGIISVVPVYI 200
QY 147 EIAYPVAVRGLGSCVQLMVVGGILLAYLAGWLEWRWLAVILGCVPSLSMLLMCFMPE 206
DB 201 EISQPNMRGALGSLVNLQSVTIGIMLAYLLGLFVNMVRLAILGILPCTVLIPGLFFIPES 260
QY 207 RFLTQHRROEAMAAALRFLWGSEQGWEDPPIGAE-----QSFHLALLRQPG 252

DB 63 FLCTLIVAGPIQFGFTGCGYSSPTQDAIRDL-----LTLSEFSVFGSLSNVGMVGA 118
QY 87 GGLVDRAGKLSLLCSVPFVAGFAVITAADQVWMLLGGRLITGLACGVASIVAPVYI 146
DB 119 SGQMAEYIGRKSLMIAAIPNIIGWLAIISFAKDSFLYMGRLLEGFGVGIISVVPVYI 178
QY 147 EIAYPVAVRGLGSCVQLMVVGGILLAYLAGWLEWRWLAVILGCVPSLSMLLMCFMPE 206
DB 179 EISQPNMRGALGSLVNLQSVTIGIMLAYLLGLFVNMVRLAILGILPCTVLIPGLFFIPES 238
QY 207 RFLTQHRROEAMAAALRFLWGSEQGWEDPPIGAE-----QSFHLALLRQPG 252
DB 239 RFLAKNMMDDEFTSLQVLRGFE-----TDISEVNDIKRAVASANKRTTIRFOELNOKK 293
QY 253 IYKPFPIIGVSLMAFOQLSGVNAVMEYAEITFEAKFKDSSLASVVGVIQVLTAVAAALIM 312
DB 294 YRPLILGILLVLOQLSGINGILFYAGSIFKAAGTNSDLATCALGAIQVLTAVGTWLV 353
QY 313 MDRAGRRLLLVSGVVMVFSTAFGAYFKLTQGGPGNSHVAISAPVSAQPVDSVGLAW 372
DB 354 LDRAGRRLLIISAGMTLSLLAVAVVFFL-KDSISQDSHMYT-----LSM 399
QY 373 LAVGSMCLFIAGFVAGWGPPIKMLMSEIFPLHVKVATGICVLTNMLMAFLVTKFSSLM 432
DB 400 ISLVALVAFVIAFSFGMGAIPIIMSEILPVSIKSLAGSEATLANMLTSFGITMT-ANLM 458
QY 433 EVLRPYGAFWLASAFICFISVLTFLFCVPETKGTLEQITAHF 474
DB 459 LWSAGGTFVSYMVVSAFTLVFVILMVVPETKGTLEBIIQFSF 500

RESULT 10
US-10-425-114-63431
; Sequence 63431, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Steven, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 63431
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4371-017-Fil_FLI.pap
US-10-425-114-63431

Query Match 28.6%; Score 702; DB 12; Length 523;
Best Local Similarity 35.9%; Pred. No. 3.2e-53;
Matches 166; Conservative 80; Mismatches 178; Indels 38; Gaps 6;

QY 27 FLAFAAALGPIISFGFALGYSSPAIPSLQRAAPPAPRLDDAAASWFGAVVTILGAAAGGV 86
DB 85 FLCTLIVAGPIQFGFTGCGYSSPTQDAIRDL-----LTLSEFSVFGSLSNVGMVGA 140
QY 87 GGLVDRAGKLSLLCSVPFVAGFAVITAADQVWMLLGGRLITGLACGVASIVAPVYI 146
DB 141 SGQMAEYIGRKSLMIAAIPNIIGWLAIISFAKDSFLYMGRLLEGFGVGIISVVPVYI 200
QY 147 EIAYPVAVRGLGSCVQLMVVGGILLAYLAGWLEWRWLAVILGCVPSLSMLLMCFMPE 206
DB 201 EISQPNMRGALGSLVNLQSVTIGIMLAYLLGLFVNMVRLAILGILPCTVLIPGLFFIPES 260
QY 207 RFLTQHRROEAMAAALRFLWGSEQGWEDPPIGAE-----QSFHLALLRQPG 252

```
Db 261 RWLAKNNMDDPETSLOVLRGFE-----TDISAENVNDIKRAVASANKRTTIRFQELNQKK 315
Qy 253 IYKPIILGVSMAFQOLSGVNAVMFYAETIFEEAKFKDSSLASVVGVIQVLFVAAL 312
Db 316 YRTPLILGILVLOQLSGINGILFYAGSIFKAAGLTNSDLATCALGAIQVLATGVTTWL 375
Qy 313 MDRAGRLLLVLSGVVMVFSTSAFGAYFKLTQGGPGNSHVAISAPVSAQPVDAASVGLAW 372
Db 376 LDRAGRLLIILISSAGMTLSLLAVAVFEL-KDSISQDSHMYIT-----LSM 421
Qy 373 LAVGSNCMLFIAGVWGPIPWLLMSHIFPLHVKGVATGICVLTNWLMAFLVTKFSSLM 432
Db 422 ISLVALVAFVIAFSGMGAIPIWIMSEILPVSIKSLAGSFATLANWLTSGFITWT-ANLM 480
Qy 433 EVLRYGAFWLASAFICISVLTFCVPETKTKLEQITAHF 474
Db 481 LWSAGGTFVSYMVVSAFTLVFVILWVPETKGRTELEIOWSF 522

RESULT 11
US-10-425-114-59933
; Sequence 59933, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59933
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-227-G11_FLI.pbp
US-10-425-114-59933

Query Match 28.6%; Score 701.5; DB 12; Length 509;
Best Local Similarity 34.2%; Pred. No. 3.4e-53;
Matches 173; Conservative 86; Mismatches 182; Indels 65; Gaps 8;

Qy 10 QPLLG-----PFGSAPRGR-----RVFLAAFA-----AALGPLSFGF 42
Db 27 KPLINTGSWYRMPAGGVMGSRQSSLMERLGSASFSLRDVAISATLCTLIIVAGPIQGF 86
Qy 43 ALGYSSPAISLQRAAPAPRLDDAASWFAVTLGAAAGVILGGMVDRAGKLSLL 102
Db 87 TCGYSSTP-----QDAITADLSLSEFSLFGLSNGVMGAMVGAISGQLAIEYIGRKSIMI 142
Qy 103 CSVPFVAGFAVITAQDVMMLGGRLLTGLACGVASLVAIPVYISEIAYPAVRGLIGSCVQ 162
Db 143 AAINIIGLWLAISAKUSSFLFMRLLGFGVGVISYTPVYIAEIPDQORGALGSVQ 202
Qy 163 LMVVVIGLLAYLAGVLEWRWLAVLGCVPSPSLLMLLMCFMPTPRFLTQHRROEAMAL 222
Db 203 LSVTIGILLAYLFGFVWRILAVLILPCSLIPGLFFVFPESPRWLAKMKMEDFEYSL 262
Qy 223 RFLWSEOGWEDPPIGAE-----QSFLALLROPGIYKPFIFIGVSLMAFQ 268
Db 263 QVLKGFQ-----TDITAEVNIKSLASSRRRTTIRFADIKQKRYSPVLVIGILVLIQ 317
Qy 269 LSGVNAVMFYAETIFEEAKFKDSSLASVVGVIQVLFVAALIMDRAGRLLLVLSGV 328
Db 318 LSGVNGILFYAASIFKAAGTNSNLATFGLGAVQVIATGVTTWLTDRAGRLLLIISTTG 377
```

```
Qy 329 MVFSTSAFGAYFKLTQGGPGNSHVAISAPVSAQPVDAASVGLAWLAVGSMCLFIAGFAVG 388
Db 378 WVTILVIVSVSFFVKDNTAAAGSHLYSV-----MSMLSLAGLWAFVFAFSLG 423
Qy 389 WGPFWLLMSHIFPLHVKGVATGICVLTNWLMAFLVTKFSSLMELVLRPYGAFWLASAF 448
Db 424 LGAIPIWIMSEILPVNIKSLAGSVATLANWLTAWAITWT-ASLMLNWSGGTFAIYAVVS 482
Qy 449 IPSVLFTLFCVPETKTKLEQITAHF 474
Db 483 TMAILFVCLWVPETKGRTELEIAFSF 508

RESULT 12
US-10-051-909-36
; Sequence 36, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 36
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (528)
; OTHER INFORMATION: Xaa = ANY AMINO ACID
US-10-051-909-36

Query Match 26.8%; Score 658.5; DB 13; Length 553;
Best Local Similarity 34.0%; Pred. No. 2.4e-49;
Matches 159; Conservative 77; Mismatches 189; Indels 43; Gaps 6;

Qy 16 PGGSAPRGR-----VFLAFAAALGPLSGFALGYSIPAIPSLQRAAPPAPRLDDAAS 70
Db 84 PGTSSMAVLRESHVSAFLCTLIIVAGPIQGFSTGFSPTQDAMVRDL---NLSISEFS 139
Qy 71 WFGAVVTILGAAAGVGLGVLDVDRAGKLSLLLCSPVPFVAGFAVITAQDVMMLGGRLLT 130
Db 140 AFGSLNVGVMGVAIGASQMAEYIGRKGLMAIPIIIGWLAIKADASFVWGRLL 199
Qy 131 GLACVASLVAIPVYISEIAYPAVRGLIGSCVQLMVVVGILLIAYLAGVLEWRWLAVLGCV 190
Db 200 GFGVGIISYTPVYIAEISQNMRRGALGSVNLSTVTFGIPLAYLLGMFIWRLAVIGAL 259
Qy 191 PPSLMLLLCMFPTPRFLTQHRROEAMALRELWSEOGWEDPPIGAE-----240
Db 260 PCTMLIPGLFFIPESPRWLAKMNLTECETSLQVLRGFE-----TDITTEVNDIKRAVAS 314
Qy 241 ----QSFLALLROPGIYKPFIFIGVSLMAFQOLSGVNAVMFYAETIFEEAKFKDSSLASV 296
Db 315 SSKEHTTISFQELNQKKYRTPLLIGILLVLIQNLISINGVILFYASSIFKAAGVTNSDLATC 374
Qy 297 VVGVIQVLFVAALIMDRAGRLLLVLSGVVMVFSTSAFGAYFKLTQGGPGNSHVAIS 356
Db 375 SLGAIQVLATGVTTWLTDRAGRLLIISTSGMTLCLLAVSVWFLLKDNISQDSNSYYI- 433
Qy 357 APVSAQPVDAASVGLAWLAVGSMCLFIAGFAVGWGPFWLLMSHIFPLHVKGVATGICVLT 416
Db 434 -----LTMISLVGIVSFVITPFGMGAIPLWMMSEILPVSIKSLGSGSIATLA 480
```

QY 417 NWMLAFVTKERSLMEVLPYCAFWLASAFCLFSVLFILFCVPTKG 464
DB 481 NWLTSPAITMT-TNMLTWSVGTFSLSYMVVSFAITVFVVLWVPTKG 527

RESULT 13

US-10-310-154-725

; Sequence 725, Application US/10310154
; Publication No. US20030233670A1

; GENERAL INFORMATION:

; APPLICANT: Edgerton, Michael D

; APPLICANT: Chomet, Paul S.

; APPLICANT: Adams, Thomas H

; APPLICANT: Ruff, Thomas G.

; APPLICANT: Agarwal, Ameeta K.

; APPLICANT: Ahrens, Jeffrey E.

; APPLICANT: Ball, James A.

; APPLICANT: Banu, G.

; APPLICANT: Bell, Erin

; APPLICANT: Boddupalli, Raghava

; APPLICANT: Deikman, Jill

; APPLICANT: Deng, Molian

; APPLICANT: Dong, Jinzhao

; APPLICANT: Duff, Stephen M.

; APPLICANT: Galligan, Meghan M.

; APPLICANT: Hinchey, Brenda S.

; APPLICANT: Huang, Shihshien

; APPLICANT: Johnson, G. Richard

; APPLICANT: Jung, Vincent

; APPLICANT: Kretzmer, Keith A

; APPLICANT: Laccetti, Lucille B.

; APPLICANT: Lai, Chao-Qiang

; APPLICANT: Lee, Gary

; APPLICANT: Lin, Jie-Yi

; APPLICANT: Liu, Jingdong

; APPLICANT: Lu, Bin

; APPLICANT: Luethy, Michael M.

; APPLICANT: Lund, Adrian

; APPLICANT: Madison, Linda L.

; APPLICANT: Malloy, Kathleen A.

; APPLICANT: McKiel, Christine L.

; APPLICANT: Miller, Philip W.

; APPLICANT: Padmavathi, Manchikanti

; APPLICANT: Parnell, Laurence D.

; APPLICANT: Start, William G.

; APPLICANT: Tennesen, Dan

; APPLICANT: Vidya, K.R.

; APPLICANT: Wang, Haiyun

; APPLICANT: Xin, Zhanqun

; APPLICANT: Xu, Nanfei

; APPLICANT: Yang, Chunzhi

; APPLICANT: Zeng, Xiaoping

; APPLICANT: Zhang, Qiang

; APPLICANT: Zhao, Yajuan

; APPLICANT: Zhou, Li

; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants

; FILE REFERENCE: 38-15(52796)B

; CURRENT APPLICATION NUMBER: US/10/310,154

; CURRENT FILING DATE: 2002-12-04

; PRIOR APPLICATION NUMBER: 60/337,358

; PRIOR FILING DATE: 2001-12-04

; NUMBER OF SEQ ID NOS: 736

; SEQ ID NO 725

; LENGTH: 481

; TYPE: PRT

; ORGANISM: Glycine max

US-10-310-154-725

Query Match 26.0%; Score 639; DB 15; Length 481;

Best Local Similarity 34.9%; Pred. No. 1e-47;

Matches 161; Conservative 76; Mismatches 182; Indels 42; Gaps 7;

QY 26 VFLAFAAALGFLSGFALGYSSPAIPSLQRAAPPAPRLDDAAASWFGAVTLGNAAGV 85

DB 45 VLLSTLVAVCGSFTFGTCVGSAPT---QAARADNLNLSAEFSWFGSLVIGAMLGA 100
QY 86 LGGLVDRAGRKLKSLLLCSVPFVAGFAVITAQDVWMLLGGRLTLTGACGVASLVPVYI 145
DB 101 TSGRTDFIGRKAMRISTGFCITGWIAVFFSKGSYSLDFGRFTGYGIGVISYVVPVYI 160
QY 146 SEIAYPAVGLGSCVOLMVVVGILLAYLAGVLEWRWLAVLGCVPSPSLMLLMLCMFET 205
DB 161 AELAPKNLRGLATTNQLLIVTGGSVFLLGSVINRELALAGLVFCICLLVGLCFIPES 220
QY 206 PRFLITQHRROEAMALREFLWSEQGWEDPP-----IGAEQSF---HLALLEPQGIYKP 256
DB 221 PRWLAKVGREKFQALSRKGDADISDEAAEILDYIETLQSLPKTKLLDLFPQSKVHS 280
QY 257 FIGVSLMAFQQLSGVNAVWFYAETIFERAKFKDSSLASVVVGVIOVLFTVAVALIMDRA 316
DB 281 VVIGVGLMACQOSVINGIGFYTAEIFVAAGLSSGKAGTIAYACIQIPFTLLGAILMDKS 340
QY 317 GRRLILVLS--GVVMVFTSAGAYFKLTQGGPGNSSHVAISAPVSAQPVDAVSGLAW-- 372
DB 341 GRPLVMWSAAGTFLGCFVAAFAFLK-----DQSLLEPFWVP 377
QY 373 -LAVGSMCLFIAGFAVGWGEIPWLLMSEIFPLHVKGVATGICVLTNWLMAFLVTKFSSL 431
DB 378 ILAFAGVLIVIAAFSIGLGSVPWVINSIFPIHLKGTAGSLVVLVAVLGAWVVSYTFNEL 437
QY 432 MEVLRYGAPWLASAFICFSVLTFLFCVPETKTKTLEQITA 472
DB 438 MSWSSP-GTLFLYAGCSLLTILFVAKLVPTKGTLEEQIA 477

RESULT 14

US-09-794-822-11

; Sequence 11, Application US/09794822

; Publication No. US20030017585A1

; GENERAL INFORMATION:

; APPLICANT: Curagen Corporation

; TITLE OF INVENTION: Novel Sodium/Solute Symporter-Like Protein and Nucleic

; FILE REFERENCE: 15966-687

; CURRENT APPLICATION NUMBER: US/09/794,822

; CURRENT FILING DATE: 2001-02-26

; PRIOR APPLICATION NUMBER: 60/185,198

; PRIOR FILING DATE: 2000-02-25

; PRIOR APPLICATION NUMBER: 60/186,410

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 11

; LENGTH: 488

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-794-822-11

Query Match 25.0%; Score 615; DB 12; Length 488;

Best Local Similarity 35.0%; Pred. No. 1.4e-45;

Matches 179; Conservative 74; Mismatches 168; Indels 90; Gaps 18;

QY 30 AFAAALGPLSGFALGYSSPAIP-----SLQRAAPPAPRLDDAAASWFGAVVT 77

DB 2 ALVAALGG---GFLFGYDTGTGIGFGLALIDFLFRFGLLTSSGALAEVLGYSTVLGLVVS 58

QY 78 ---LGAAGGVGLGWLVDRAGKLSLLCSVPFVAGFAVITAQ-----DVW---MLGG 126

DB 59 IFFLRLIGSLFAGKLGDRFRKKSLLIALVLVFIGALLSGAAGPYTTIGLWAFYLLIVG 118

QY 127 RLITGLACGVASLVAPVYISETIAYPAVRGLLGSVCQVLMVVVGIILAYLAGVWL----- 179

DB 119 RVLVGLGVGASGLVMPVYISETIAPKALRGALGSYQAITIGILVVAIIGLGNKTNDS 178

QY 180 ---EWRMLAVLG--CVPPSILMLLMCMFETPFLTQHRROEAMALRFLWSEQ--GWE 233

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 19:24:47 : Search time 34 Seconds
(without alignments)
724.282 Million cell updates/sec

Title: US-09-886-954A-1

Perfect score: 2457

Sequence: 1 MTEDPRTPTQLLPFGGSA.....CVPTKTKTLEQITAHFEGR 477

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	615	25.0	488	4	US-10-162-012-46
2	605	24.6	514	4	US-09-489-039A-11902
3	578.5	23.5	476	4	US-09-489-039A-11933
4	575.5	23.4	494	2	US-09-031-392-5
5	575.5	23.4	494	3	US-09-299-549-5
6	575.5	23.4	494	4	US-09-610-417-5
7	568	23.1	493	2	US-09-031-392-10
8	568	23.1	493	3	US-09-299-549-10
9	568	23.1	493	4	US-09-610-417-10
10	548.5	22.3	518	4	US-09-679-686B-23
11	542.5	22.1	517	4	US-09-679-686B-18
12	536	21.8	509	2	US-09-031-392-6
13	536	21.8	509	3	US-09-299-549-6
14	536	21.8	509	4	US-09-610-417-6
15	535	21.8	514	4	US-09-679-686B-22
16	529.5	21.6	502	4	US-09-679-686B-2
17	526.5	21.4	519	4	US-09-679-686B-24
18	516.5	21.0	518	4	US-09-679-686B-16
19	516	21.0	584	2	US-08-928-692-13
20	516	21.0	584	4	US-09-339-972-13
21	513	20.9	524	2	US-08-928-692-12
22	513	20.9	524	4	US-09-339-972-12
23	507.5	20.7	501	4	US-09-489-039A-11731
24	506.5	20.6	513	4	US-09-291-922-20
25	498.5	20.3	492	2	US-08-355-844-3
26	498.5	20.3	492	5	PCT-US95-16126-3
27	497	20.2	549	4	US-09-291-922-30

ALIGNMENTS

RESULT 1

US-10-162-012-46

; Sequence 46, Application US/10162012

; Patent No. 6682597

; GENERAL INFORMATION:

; APPLICANT: Curtis, Rory A.J.

; APPLICANT: Silos-Santiago, Inmaculada

; APPLICANT: Gu, Wei

; TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS

; FILE REFERENCE: 10448-190001

; CURRENT APPLICATION NUMBER: US/10/162.012

; CURRENT FILING DATE: 2002-06-04

; PRIOR APPLICATION NUMBER: US 60/209,845

; PRIOR FILING DATE: 2000-06-06

; PRIOR APPLICATION NUMBER: US 09/875,321

; PRIOR FILING DATE: 2001-06-06

; PRIOR APPLICATION NUMBER: PCT/US01/18340

; PRIOR FILING DATE: 2001-06-06

; PRIOR APPLICATION NUMBER: US 60/209,257

; PRIOR FILING DATE: 2000-06-05

; PRIOR APPLICATION NUMBER: US 09/875,423

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: PCT/US01/18398

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/209,238

; PRIOR FILING DATE: 2000-06-05

; PRIOR APPLICATION NUMBER: US 09/875,363

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: PCT/US01/18247

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/227,068

; PRIOR FILING DATE: 2000-08-22

; PRIOR APPLICATION NUMBER: US 09/928,530

; PRIOR FILING DATE: 2001-08-13

; PRIOR APPLICATION NUMBER: PCT/US01/25475

; PRIOR FILING DATE: 2001-08-15

; PRIOR APPLICATION NUMBER: US 60/226,770

; PRIOR FILING DATE: 2000-08-21

; PRIOR APPLICATION NUMBER: US 09/934,421

; PRIOR FILING DATE: 2001-08-21

; PRIOR APPLICATION NUMBER: PCT/US01/26096

; PRIOR FILING DATE: 2001-08-21

; PRIOR APPLICATION NUMBER: US 60/279,281

; PRIOR FILING DATE: 2001-03-28

; PRIOR APPLICATION NUMBER: US 10/109,029

; PRIOR FILING DATE: 2002-03-28

; PRIOR APPLICATION NUMBER: PCT/US02/09728

; PRIOR FILING DATE: 2002-03-28

; PRIOR APPLICATION NUMBER: US 60/290,288

; PRIOR FILING DATE: 2001-05-11

Sequence 9, Appli
Sequence 9549, Ap
Sequence 4, Appli
Sequence 4, Appli
Sequence 19, Appl
Sequence 10, Appl
Sequence 10, Appl
Sequence 44, Appl
Sequence 26, Appl
Sequence 24, Appl
Sequence 28, Appl
Sequence 3, Appli
Sequence 3, Appli
Sequence 22, Appl
Sequence 12, Appl
Sequence 21, Appl


```
Query Match      23.4%; Score 575.5; DB 3; Length 494;
Best Local Similarity 30.7%; Pred. No. 5.3e-46;
Matches 150; Conservative 94; Mismatches 179; Indels 65; Gaps 10;

QY 26 VFLAFAAALGSLFSGFALGYSSPAIPSL-----QRAAPPAPRLDDAAAASWFGAV 75
Db 10 LIPALSIATIGSFQGYNTGVINAPEAIKDFLNTLEERSETPPSSVLLTSLWSLSVAI 69
QY 76 VTGAAAGGVLGWLVDRAGRKLSLLGCVFPVAGFAVI---TAAQDVWMLGGRLTLGL 132
Db 70 FSVGGMIGSFSVGLFVNRFGRRNSMLIVNLLAIAAGCLMGFCCKIAESVEMILGLIIGL 129
QY 133 AGVASLAVPVYISIAIYPAVRGLGSCVQLMVVVGILLAYLAGWVL---EWRMLAVLG 188
Db 130 FGLCTGFVPYVIGISPTALRGAGTINQLGIVIGILVAQIFGLKVLGTEDLWPLLLIG 189
QY 189 --CVPSPMLLMCFMPTPRFLTQHRROE-AMAAALFLMGSEOGWEDPPIGASQSFH- 244
Db 190 FTILPAIIQCAALPCPESPREFLLINRKEEKAKELQRLWCTEDVAQDIQEMKDEMRM 249
QY 245 -----LALLRQFGIYKPFIIIGVSLMAFOQLSGVNAVVFYAEIIFEEAKFKDSSLASV 296
Db 250 SOEKQVTVLELFRAPNYRQPIIISIMLQLSQQLSGINAVFYSTGIFKDGAVQEPVYATI 309
QY 297 VVGVIQVLTAAALIMDRAGRRLLLVLSGVVMVFTSAFGAYFKLTGGGPNSSHVAIS 356
Db 310 GAGVNTIFTVSVFLVERAGRTLHLI-----GLGMAFCSIL 348
QY 357 APVSAQPDVDSVGLAWLAVGSMCLFIAGFVAGWGPIWMLMSEIFPLHVKGVATGICVLT 416
Db 349 MTISILLKDNYSWMSFCIGAILVFVAFEGPIPWFIWAELEFGQGPRAAMAVAGCS 408
QY 417 NWLMAFLVTKFESSLMELVRPGAPWL-ASAFICFSV-----LFTLCVBPETKGTLEQ 469
Db 409 NWTSNFLV-----GLLFPSSATFYLGAYVIVTFVFLVFWVTFPKVPETRGRTPEE 460

QY 470 ITAHFEGR 477
Db 461 ITRAPEGQ 468

RESULT 7
US-09-031-392-10
; Sequence 10, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:

QY 26 VFLAFAAALGSLFSGFALGYSSPAIPSL-----QRAAPPAPRLDDAAAASWFGAV 75
Db 10 LIPALSIATIGSFQGYNTGVINAPEAIKDFLNTLEERSETPPSSVLLTSLWSLSVAI 69
QY 76 VTGAAAGGVLGWLVDRAGRKLSLLGCVFPVAGFAVI---TAAQDVWMLGGRLTLGL 132
Db 70 FSVGGMIGSFSVGLFVNRFGRRNSMLIVNLLAIAAGCLMGFCCKIAESVEMILGLIIGL 129
QY 133 AGVASLAVPVYISIAIYPAVRGLGSCVQLMVVVGILLAYLAGWVL---EWRMLAVLG 188
Db 130 FGLCTGFVPYVIGISPTALRGAGTINQLGIVIGILVAQIFGLKVLGTEDLWPLLLIG 189
QY 189 --CVPSPMLLMCFMPTPRFLTQHRROE-AMAAALFLMGSEOGWEDPPIGASQSFH- 244
Db 190 FTILPAIIQCAALPCPESPREFLLINRKEEKAKELQRLWCTEDVAQDIQEMKDEMRM 249
QY 245 -----LALLRQFGIYKPFIIIGVSLMAFOQLSGVNAVVFYAEIIFEEAKFKDSSLASV 296
Db 250 SOEKQVTVLELFRAPNYRQPIIISIMLQLSQQLSGINAVFYSTGIFKDGAVQEPVYATI 309
QY 297 VVGVIQVLTAAALIMDRAGRRLLLVLSGVVMVFTSAFGAYFKLTGGGPNSSHVAIS 356
Db 310 GAGVNTIFTVSVFLVERAGRTLHLI-----GLGMAFCSIL 348
QY 357 APVSAQPDVDSVGLAWLAVGSMCLFIAGFVAGWGPIWMLMSEIFPLHVKGVATGICVLT 416
Db 349 MTISILLKDNYSWMSFCIGAILVFVAFEGPIPWFIWAELEFGQGPRAAMAVAGCS 408
QY 417 NWLMAFLVTKFESSLMELVRPGAPWL-ASAFICFSV-----LFTLCVBPETKGTLEQ 469
Db 409 NWTSNFLV-----GLLFPSSATFYLGAYVIVTFVFLVFWVTFPKVPETRGRTPEE 460

QY 470 ITAHFEGR 477
Db 461 ITRAPEGQ 468

RESULT 6
US-09-610-417-5
; Sequence 5, Application US/09610417
; Patent No. 6346374
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/610,417
; FILING DATE: 05-JUL-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/299,549
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
```


Db 455 ETKGRTFDEIAAAF 468

RESULT 9
US-09-610-417-10
; Sequence 10, Application US/09610417
; Patent No. 6346374
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/610,417
; FILING DATE: 05-Jul-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/299,549
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; LENGTH: 493 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-610-417-10

Query Match 23.1%; Score 568; DB 4; Length 493;
Best Local Similarity 32.2%; Pred. No. 2.7e-45;
Matches 159; Conservative 87; Mismatches 160; Indels 88; Gaps 14;

Qy 28 LAFAAALGPLSGFALGYSSPAIPSLQRAAPPAPRLDDAAASWFG-----AVVTILGA- 80
Db 16 LAVLIAALGSPQYGNLGVIN-----APOKVEAFYETWLGRGXGXPSTLTLL 65
Qy 81 -----AAGVLGWLIV-----DRAGKLSLLCSVPFVAG---FAVITAAQDVMMLG 125
Db 66 WLSVSIFAVGGMIGSFVIGXIGNRLGKXXAMLVNNVLAIGLLMGLAKXAXSFEMLIL 125
Qy 126 GRLTGLACGVASIVAPVYISEIAYPAVRGLLGSVCQVMVVVHLLAYLAGWVLE----- 180
Db 126 GRFTIGLYCGLSSGWPVMYGEISFTALRGALGTNLQGLVIGILIAQVLG--LDSLLGN 183
Qy 181 ---NRWLAVLGVCPESLMLLMCFMPTERTFLLTQHRROEAMA--ARFLWG----- 227
Db 184 ESLWPLLLGLTGVALLQLLLPCEPSPRLLI--KNNEEARAKALQRURGTADVSOEV 242
Qy 228 SEQGWEDPPIGAQSFH--LALLRQPIYKPIIFGVSLMAFOQLSGVNAVMPFYAETIFEAA 286
Db 243 AEMKDESRMXKSEKXSVLELFRSXYRQPVIITAIVLQSLQSLGINAVFYYSISPEKA 302
Qy 287 KFKDSSLASVVVGVIVQLVFTAAALIMDRAGRRLLLVLGVVWVFSTSAFGAYPKLTQGG 346

Db 303 GVGQPVYATIGAGVNTVFTVSVFVERAGRTLELL-----G 341
Qy 347 PGNSSHVAISAPVSAQPDVASVGLAWLAVGSMCLFIAGFAVGWGPIMPMLLSEIFPLHVK 406
Db 342 LGMAGCAVLTIALALLDQVPMWSYVSIIVAFGEVAFVGPBPFWFIVAELEPSQGR 401
Qy 407 GVATGICVLTNWLMAFLVTKSFSSLMELVRPYGAFWLASAFCISVL-----FTLCVP 460
Db 402 PAATAVAGFSNWTSNFIVGLLFQYIAELLGPY-----VFIVFAVLLLLFIETFLKVP 454
Qy 461 ETKGKTLQITAHF 474
Db 455 ETKGRTFDEIAAAF 468

RESULT 10
US-09-679-686B-23
; Sequence 23, Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
; FILE REFERENCE: BB160 US NA
; CURRENT APPLICATION NUMBER: US/09/679,686B
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/081,131
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: PCT/US99/07561
; PRIOR FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Medicago truncatula
US-09-679-686B-23

Query Match 22.3%; Score 548.5; DB 4; Length 518;
Best Local Similarity 30.4%; Pred. No. 2e-43;
Matches 151; Conservative 88; Mismatches 190; Indels 57; Gaps 16;

Qy 16 PGGSAPRRVRVFLAFAAALGPLSGFALGYSSPAI---PSLQRAAPPAPR----- 63
Db 16 PGNLTP---FVTITCIVAAAGGLIFGYDIGISGVTSMDPPLKFFFAVYKKNKDKSTN 72
Qy 64 ---LDDAASWFGAVVTILGAAGVGLGGLVDRAGKLSLLCSVPFVAGFAVITAAQD 119
Db 73 QYCYQDSTLMTFTSSLYLAALLSSLVASTITRFRGRKLSMLFGLLFLVGLALINGFANH 132
Qy 120 VMMLLGRLLTGLACGVASIVAPVYISEIAYPAVRGLLGSVCQVMVVVHLLAYLAGWVL 179
Db 133 VMMLIVGRILLGFGIGFANQPVPLYSEMAYPKYRGALNIGFQISITIGILVANVLYFF 192
Qy 180 E-----WRW--LAVLGVCPSPSLMLLL--MCFMPTERTFLLTQHRROEAMAARFLWSE--- 229
Db 193 AKIKGSGWGLSLGGAMVPALITIGSLVLPDTPNSMIERGDRDGAQAKLRKIRGIEDVD 252
Qy 230 QGWEDPPIGAQSFHIA-----LRLQPIYKPF--FICVSLMAFOQLSGVNAVMPFYAETIF 283
Db 253 EEFNDLVAASEASQVENPWRNLLQK--YRPQUTMAVLIPFQOFTGINVIMYAPVLF 310
Qy 284 EEAKFKD--SSLASVVVGVIVQLVFTAAALIMDRAGRRLLLVLGVVW---VFSTSAFG 337
Db 311 NSIGFKDDASLMSAVITGVVNVVATCVSIYGVDKWGRRALFLEGAQMLICQAVAAAG 370
Qy 338 AYFKLTQGGGNSHVAISAPVSAQPDVASVGLAWL---VGSMLCFIAGFAVGWGPFW 394
Db 371 AKFG-TSGNPNLPE-----WYAIVVVVLFTICIVAGFAWSGWLPGW 410

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 19:08:41 ; Search time 126 Seconds
(without alignments)
1069.644 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTFEDPBTQPLLPQGGSA.....CVPTKTKLTLEQITAHPEGR 477

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB ID	Description
1	2457	100.0	477	4	Aae04888 Human tra
2	2457	100.0	477	6	Abp58364 Human sol
3	2448	99.6	477	4	Aab66932 Human GLU
4	2224.5	90.5	478	4	Aab66939 GLUTX1 co
5	2162.5	88.0	478	4	Aab66933 Rat GLUTX
6	2150	87.5	477	4	Aab66934 Murine GL
7	1491	60.7	326	5	Abb89717 Human pol
8	1135	46.2	262	4	Aae06579 Human pro
9	1131	46.0	248	5	Abu11283 cDNA enco
10	969.5	39.5	503	4	Aab66938 Rat GLUTX
11	953	38.8	507	4	Aab66941 GLUTX3 co
12	948	38.6	507	4	Aab66937 Human GLU
13	948	38.6	507	4	Aam93417 Human pol
14	948	38.6	507	6	Ada84077 Human SLC
15	948	38.6	507	6	Abp58365 Human sol
16	948	38.6	507	6	Abp58363 Human sol
17	894	36.4	445	4	Aam93364 Human pol
18	788.5	32.1	489	4	Abb63511 Drosophil
19	788	32.1	539	4	Abb65350 Drosophil
20	764.5	31.1	433	4	Abb63758 Drosophil
21	727	29.6	465	4	Abb62408 Drosophil
22	705	28.7	497	4	Abb57922 Drosophil
23	670	27.3	447	3	Aag13749 Arabidops
24	670	27.3	463	3	Aag13748 Arabidops
25	669.5	27.2	437	3	Aag13750 Arabidops

ALIGNMENTS

RESULT 1

AAE04888
ID AAE04888 standard; protein; 477 AA.

AC AAE04888;

DT 10-SEP-2001 (first entry)

XX Human transporter and ion channel-1 (TRICH-1) protein.

XX Human; transporter and ion channel-1; TRICH-1; vaccine; cystic fibrosis;
KW gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy;
KW hypertension; angina; neurological disorder; asthma; bipolar disorder;
KW dementia; depression; Alzheimer's disease; epilepsy; mood; arrhythmia;
KW Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke;
KW Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy;
KW demyelinating disease; mental disorder; Schizophrenia; polymyositis;
KW muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease;
KW dermatomyositis; diabetes mellitus; immunological disorder; psoriasis;
KW rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus;
KW sickle cell anaemia; Wilson's disease; infertility; Cushing's disease;
KW scleroderma; pulmonary artery stenosis; neutropenic; Addison's disease;
KW malabsorption syndrome; hypercholesterolaemia; cancer.

XX Homo sapiens.

Key	Location/Qualifiers
Domain	/note= "Sugar transporter domain"
Domain	29..474
Domain	259..279
Domain	/label= Transmembrane_domain
Domain	293..313
Domain	/label= Transmembrane_domain
Domain	320..339
Domain	/label= Transmembrane_domain
Domain	438..457
Domain	/label= Transmembrane_domain

WO200146258-A2.

28-JUN-2001.

22-DEC-2000; 2000WO-US035095.

23-DEC-1999; 99US-0172000P.

14-JAN-2000; 2000US-0176083P.

21-JAN-2000; 2000US-0177332P.

28-JAN-2000; 2000US-0178572P.

CC The present sequence is that of human solute carrier type 2A (SLC2A)
 CC polypeptides 12735153 and 7657681. In the present invention, genetic
 CC screens were designed to identify modifiers of the p53 pathway in
 CC Drosophila in which p53 was overexpressed. Human orthologues
 CC (polynucleotides and polypeptides) of one such modifier were then
 CC identified, including the present polypeptides. SLC2As are glucose
 CC transporter proteins with sugar transporter domains. They are attractive
 CC drug targets for treatment of pathologies associated with a defective p53
 CC signalling pathway, such as cancer. The invention provides in vitro and
 CC in vivo methods of assessing SLC2A function. Modulation of an SLC2A or
 CC its binding partners is useful for understanding the association of the
 CC p53 pathway and its members in normal and disease conditions and for
 CC developing diagnostic and therapeutic modalities of p53-related
 CC pathologies. SLC2A-modulating agents that act by inhibiting or enhancing
 CC SLC2A expression, directly or indirectly, e.g. by affecting an SLC2A
 CC function such as transport or binding activity, can be identified using
 CC methods provided. Modulators include small molecules, nucleic acids,
 CC antibodies, antisense oligonucleotides and phosphothioate morpholino
 CC oligomers (claimed)

XX Sequence 477 AA;

Query Match 100.0%; Score 2457; DB 6; Length 477;
 Best Local Similarity 100.0%; Pred. No. 5.1e-227;
 Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTPDPETQLLPGPGSAPGRVFLAFAAALGPLSGFALGYSSPAIPSLQRAAPP 60
 Db 1 MTPDPETQLLPGPGSAPGRVFLAFAAALGPLSGFALGYSSPAIPSLQRAAPP 60
 Qy 61 APRLDDAAASWFGAVVTILGAAAGVGLGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDV 120
 Db 61 APRLDDAAASWFGAVVTILGAAAGVGLGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDV 120
 Qy 121 WMLLGGRLLTGLACGVASLVAPVYISIAYPVAVRGLLSCVQLMVVVGILLAYLAGWYLE 180
 Db 121 WMLLGGRLLTGLACGVASLVAPVYISIAYPVAVRGLLSCVQLMVVVGILLAYLAGWYLE 180
 Qy 181 WRWLAVLGCVPSPSLMLLLCMFETPRFLLTQHRROEAMALRFLWGSEQWEDPPIGAE 240
 Db 181 WRWLAVLGCVPSPSLMLLLCMFETPRFLLTQHRROEAMALRFLWGSEQWEDPPIGAE 240
 Qy 241 QSFHALLRQPGIYKPIIIGVSLMAFQOLSGVNAVMFYAETIFEEAKFKDSSLASVVGV 300
 Db 241 QSFHALLRQPGIYKPIIIGVSLMAFQOLSGVNAVMFYAETIFEEAKFKDSSLASVVGV 300
 Qy 301 IQVLTAAALIMDRAGRRLLLVLSGVVMVFTSAFGAYFKLTQGGPGNSHVAISAPVS 360
 Db 301 IQVLTAAALIMDRAGRRLLLVLSGVVMVFTSAFGAYFKLTQGGPGNSHVAISAPVS 360
 Qy 361 AQPVDASVGLAWLAVGSMCLPIAGFVAGWGPIPWLLMSEIFPLHVKGATGICVLTNWL 420
 Db 361 AQPVDASVGLAWLAVGSMCLPIAGFVAGWGPIPWLLMSEIFPLHVKGATGICVLTNWL 420
 Qy 421 AFLVTKFSSLMELVRPYGAFWLASAFICISVLTFLFCVPTKGTLEQITAHFEGR 477
 Db 421 AFLVTKFSSLMELVRPYGAFWLASAFICISVLTFLFCVPTKGTLEQITAHFEGR 477

RESULT 3

AAB66932 standard; protein; 477 AA.

XX AAB66932;

AC AAB66932;

XX 17-APR-2001 (first entry)

XX Human GLUTX1.

XX Human; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.

OS Homo sapiens.
 XX WO200104145-A2.
 PN 18-JAN-2001.
 XX 14-JUL-2000; 2000WO-IB001042.
 XX 14-JUL-1999; 99US-0143907P.
 PR 27-AUG-1999; 99US-0151140P.
 PR 23-FEB-2000; 2000US-0184285P.
 PR 13-JUL-2000; 2000US-00616132.
 XX (UCLA-) UNIV LAUSANNE.
 PA Thorens B, Ibberson M, Uldry M;
 PI WPI; 2001-112615/12.
 XX N-PSDB; AAF55865.
 DR Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
 XX prevention, diagnosis and treatment of hexose transport disorders, e.g.
 PT ischaemia and diabetes.
 PT Claim 11; Page 70-71; 124pp; English.
 PS The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 XX AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is human GLUTX1

XX Sequence 477 AA;

Query Match 99.6%; Score 2448; DB 4; Length 477;
 Best Local Similarity 99.6%; Pred. No. 3.7e-226;
 Matches 475; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MTPDPETQLLPGPGSAPGRVFLAFAAALGPLSGFALGYSSPAIPSLQRAAPP 60
 Db 1 MTPDPETQLLPGPGSAPGRVFLAFAAALGPLSGFALGYSSPAIPSLQRAAPP 60
 Qy 61 APRLDDAAASWFGAVVTILGAAAGVGLGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDV 120
 Db 61 APRLDDAAASWFGAVVTILGAAAGVGLGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDV 120
 Qy 121 WMLLGGRLLTGLACGVASLVAPVYISIAYPVAVRGLLSCVQLMVVVGILLAYLAGWYLE 180
 Db 121 WMLLGGRLLTGLACGVASLVAPVYISIAYPVAVRGLLSCVQLMVVVGILLAYLAGWYLE 180
 Qy 181 WRWLAVLGCVPSPSLMLLLCMFETPRFLLTQHRROEAMALRFLWGSEQWEDPPIGAE 240
 Db 181 WRWLAVLGCVPSPSLMLLLCMFETPRFLLTQHRROEAMALRFLWGSEQWEDPPIGAE 240
 Qy 241 QSFHALLRQPGIYKPIIIGVSLMAFQOLSGVNAVMFYAETIFEEAKFKDSSLASVVGV 300
 Db 241 QSFHALLRQPGIYKPIIIGVSLMAFQOLSGVNAVMFYAETIFEEAKFKDSSLASVVGV 300
 Qy 301 IQVLTAAALIMDRAGRRLLLVLSGVVMVFTSAFGAYFKLTQGGPGNSHVAISAPVS 360
 Db 301 IQVLTAAALIMDRAGRRLLLVLSGVVMVFTSAFGAYFKLTQGGPGNSHVAISAPVS 360
 Qy 361 AQPVDASVGLAWLAVGSMCLPIAGFVAGWGPIPWLLMSEIFPLHVKGATGICVLTNWL 420
 Db 361 AQPVDASVGLAWLAVGSMCLPIAGFVAGWGPIPWLLMSEIFPLHVKGATGICVLTNWL 420
 Qy 421 AFLVTKFSSLMELVRPYGAFWLASAFICISVLTFLFCVPTKGTLEQITAHFEGR 477
 Db 421 AFLVTKFSSLMELVRPYGAFWLASAFICISVLTFLFCVPTKGTLEQITAHFEGR 477

```

RESULT 4
ID AAB66939 standard; protein; 478 AA.
XX
AC AAB66939;
XX
DT 17-APR-2001 (first entry)
XX
GLUTX1 consensus sequence.
DE
DE
KW GLUTX; gene therapy; vaccine; hexose transport modulator; human; rat;
KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia; murine;
KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS Homo sapiens.
OS Mus sp.
OS Rattus sp.
XX
PN WO200104145-A2.
XX
PD 18-JAN-2001.
XX
PF 14-JUL-2000; 2000WO-IB001042.
XX
PR 14-JUL-1999; 99US-0143907P.
PR 27-AUG-1999; 99US-0151140P.
PR 23-FEB-2000; 2000US-0184285P.
PR 13-JUL-2000; 2000US-00616132.
XX
XX (UYLA-) UNIV LAUSANNE.
XX
XX Thorens B, Ibberson M, Uldry M;
XX WPI; 2001-112615/12.
XX
XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
XX prevention, diagnosis and treatment of hexose transport disorders, e.g.
XX ischemia and diabetes.
XX
XX Claim 11; Page 74-75; 124pp; English.
XX
XX The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
XX AAB66932-AAB66941). The GLUTX proteins are related to the facultative
XX glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
XX function. The GLUTX proteins may be used in the diagnosis, prevention and
XX treatment of hexose transport disorders such as ischaemia, diabetes,
XX hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
XX neurodegenerative disease. The present sequence is a consensus sequence
XX for GLUTX1
XX
XX Sequence 478 AA;
Query Match 90.5%; Score 2224.5; DB 4; Length 478;
Best Local Similarity 88.9%; Pred. No. 1.1e-204;
Matches 425; Conservative 23; Mismatches 29; Indels 1; Gaps 1;
QY 1 MTPEDPETOPLPGGSGAPRGRVFLAFAAALGPLSGFALGYSSPAIPSLQRAAPP 60
DB 1 MSPDPQETOPLRPPGAPRGRVFLAFAAALGPLSGFALGYSSPAIPSLRRTAPP 60
QY 61 APRLDDAAASFAGVAVTIGAAAGVGLGWLVDRAGRKLSLLCSVFFVAGFAVITAAQDV 120
DB 61 ALRLGDAASAFAGVAVTIGAAAGVGLGWLVDRAGRKLSLLCTVPFVTGFAVITAARDV 120
QY 121 WMLLGGRLTLGLACVASLAPVYISFIAYPAVRGLGSCVQLMVVVGILLAYLAGWLE 180
DB 121 WMLLGGRLTLGLACVASLAPVYISFIAYPAVRGLGSCVQLMVVVGILLAYLAGWLE 180
QY 181 WRWLAVLCVPPSLMLLMCFMPETPRFLTQHRREQEAMAAALRFLWGSEQGWEDPPPIGAE 240
DB 181 WRWLAVLCVPPSLMLLMCFMPETPRFLTQHRREQEAMAAALRFLWGSEQGWEDPPVCAE 240
QY 241 -QSFHALLRQPGIYKPLIIGLSLMVFOQLSGVNAIMFYAETIFEEAKFKDSSLASVVG 299

```

```

Db 241 HQGFOLALLRRPGIYKPLIIGLSLMVFOQLSGVNAIMFYAETIFEEAKFKDSSLASVVG 300
QY 300 VIQVLFTAAALIMDRAGRRLLLVLSGVVMVFSTSAFGAYFKLTQGGPNSHVAISAPV 359
Db 301 IIQVLFTAAALIMDRAGRRLLLVLSGVIMVFMSAFGTGFYFKLTQSGPNSHVGILVPI 360
QY 360 SAQPYDASVGLAWLAVGSMCLFIAGFAVGWGPPIWLLMSEIFPLHVKGVATGICVLTNWL 419
Db 361 SAEPVDVSVGLAWLAVGSMCLFIAGFAVGWGPPIWLLMSEIFPLHVKGVATGICVLTNWF 420
QY 420 MAFIVTKFSSIMEVLRPYGAFWLASAFICFISVLFVFCVPETKGTLEQITAHFEGR 477
Db 421 MAFIVTKFNSIMEVLRPYGAFWLTAFAFCILSVLFTLCVPETKGTLEQITAHFEGR 478

```

```

RESULT 5
ID AAB66933 standard; protein; 478 AA.
XX
AC AAB66933;
XX
DT 17-APR-2001 (first entry)
XX
Rat GLUTX1.
DE
DE
KW Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;
KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS Rattus sp.
XX
PN WO200104145-A2.
XX
PD 18-JAN-2001.
XX
PF 14-JUL-2000; 2000WO-IB001042.
XX
PR 14-JUL-1999; 99US-0143907P.
PR 27-AUG-1999; 99US-0151140P.
PR 23-FEB-2000; 2000US-0184285P.
PR 13-JUL-2000; 2000US-00616132.
XX
XX (UYLA-) UNIV LAUSANNE.
XX
XX Thorens B, Ibberson M, Uldry M;
XX WPI; 2001-112615/12.
XX
XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
XX prevention, diagnosis and treatment of hexose transport disorders, e.g.
XX ischemia and diabetes.
XX
XX Claim 11; Page 71-72; 124pp; English.
XX
XX The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
XX AAB66932-AAB66941). The GLUTX proteins are related to the facultative
XX glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
XX function. The GLUTX proteins may be used in the diagnosis, prevention and
XX treatment of hexose transport disorders such as ischaemia, diabetes,
XX hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
XX neurodegenerative disease. The present sequence is rat GLUTX1
XX
XX Sequence 478 AA;
Query Match 88.0%; Score 2162.5; DB 4; Length 478;
Best Local Similarity 85.4%; Pred. No. 9.9e-199;
Matches 408; Conservative 33; Mismatches 36; Indels 1; Gaps 1;
QY 1 MTPEDPETOPLPGGSGAPRGRVFLAFAAALGPLSGFALGYSSPAIPSLQRAAPP 60
DB 1 MSPDPQETOPLRPPGAPRGRVFLAFAAALGPLSGFALGYSSPAIPSLRRTAPP 60

```

QY 61 APRLDAAASWFGAVVTLLGAAAGVGLGWLVDRAGKLSLLLSVPPVAGFAVITAADV 120
 DB 61 ALRLGDTAAASWFGAVVTLLGAAAGVGLGWLVDRAGKLSLLLSCTVPPVTFGFAVITAARDV 120
 QY 121 WMLLGGRLLTGLACGVASLVAPVVISIAYPAVRGLLGSCVOLMVVVGILLAYLAGWVLE 180
 DB 121 WMLLGGRLLTGLACGVASLVAPVVISIAYPAVRGLLGSCVOLMVVVGILLAYLAGWVLE 180
 QY 181 WRWLAVLGCVPVPSLMLLLCMFMPETPRFLLTQHRROQEAALRFLMGSEGWEDPPIGAE 240
 DB 181 WRWLAVLGCVPVPTLMLLLCMYMPETPRFLLTQHQYQEAALRFLMGSEGWEEPPVGAE 240
 QY 241 -QSFHALLPQPGYKFIIGVSLMAFOQLSGVNAVVFYAEITPEBAKFKDSSLASVVVG 299
 DB 241 HQGFQALMLRRPGYKFLIIGICLMVFQQLSGVNAVFIYAEITPEBAKFKDSSLASVTVG 300
 QY 300 VIQVLTFAVAALIMDRAGRELLLVLSGVVMVFSTSAFGAYFKLTQGGPGNSHVAISAPV 359
 DB 301 IIQVLTFAVAALIMDRAGRELLLVLSGVVMVFSTSAFGAYFKLTQGGPGNSHVGILLVPI 360
 QY 360 SAQFVDASVGLAVGSMCLFIAGFAVGWGPVPLWLMSEIFPLHVKGATGICVLTNWL 419
 DB 361 SAEFADVHLGLAVLAVGSMCLFIAGFAVGWGPVPLWLMSEIFPLHVKGATGICVLTNWF 420
 QY 420 MAFIVTKFSSLMELVRPYGAFWLASAFICFVSFLTLFCVPEPKGKTLQITAHFEGR 477
 DB 421 MAFIVTKFNSVMEILRPYGAFWLTAAFCLSVLTLTVPETKGRTEQVTAHFEGR 478

RESULT 6

AA866934
 ID AA866934 standard; protein; 477 AA.

AC AA866934;

DT 17-APR-2001 (first entry)

DE Murine GLUTX1.

XX Murine; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.

XX Mus sp.

XX WO200104145-A2.

XX 18-JAN-2001.

XX 14-JUL-2000; 2000WO-IB001042.

XX 14-JUL-1999; 99US-0143907P.

XX 27-AUG-1999; 99US-01511140P.

XX 23-FEB-2000; 2000US-0184285P.

XX 13-JUL-2000; 2000US-00616132.

XX (UYLA-) UNIV LAUSANNE.

XX Thorens B, Ibberson M, Uldry M;

XX WPI; 2001-112615/12.

XX N-FSDB; AAF55867.

XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
 PT prevention, diagnosis and treatment of hexose transport disorders, e.g.
 PT ischemia and diabetes.

XX Claim 11; Page 73-74; 124pp; English.

XX The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport

CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is murine GLUTX1

XX SQ Sequence 477 AA;

Query Match 87.5%; Score 2150; DB 4; Length 477;

Best local Similarity 85.6%; Pred. No. 1.6e-197;

Matches 409; Conservative 33; Mismatches 34; Indels 2; Gaps 2;

QY 1 MTPEDPEETOPLLPGPGSAPGRRRVFLAAFAAALGPLSEFGALGVSPAIPIQRAAPP 60

DB 1 MSPEDEQETQPLLRPEARTPRGRVFLASFAAALGPLNFGALGVSPAIPISLRRAPP 60

QY 61 APRLDAAASWFGAVVTLLGAAAGVGLGWLVDRAGKLSLLLSVPPVAGFAVITAADV 120

DB 61 ALRLGDTAAASWFGAVVTLLGAAAGVGLGWLVDRAGKLSLLLSCTVPPVTFGFAVITAARDV 120

QY 121 WMLLGGRLLTGLACGVASLVAPVVISIAYPAVRGLLGSCVOLMVVVGILLAYLAGWVLE 180

DB 121 WMLLGGRLLTGLACGVASLVAPVVISIAYPAVRGLLGSCVOLMVVVGILLAYLAGWVLE 180

QY 181 WRWLAVLGCVPVPSLMLLLCMFMPETPRFLLTQHRROQEAALRFLMGSEGWEDPPIGAE 240

DB 181 WRWLAVLGCVPVPTLMLLLCMYMPETPRFLLTQHQYQEAALRFLMGSEGWEEPPVGAE 240

QY 241 -QSFHALLPQPGYKFIIGVSLMAFOQLSGVNAVVFYAEITPEBAKFKDSSLASVVVG 299

DB 241 HQGFQALMLRRPGYKFLIIGICLMVFQQLSGVNAVFIYAEITPEBAKFKDSSLASVTVG 300

QY 300 VIQVLTFAVAALIMDRAGRELLLVLSGVVMVFSTSAFGAYFKLTQGGPGNSHVAISAPV 359

DB 301 IIQVLTFAVAALIMDRAGRELLLVLSGVVMVFSTSAFGAYFKLTQGGPGNSHVGILLVPI 359

QY 360 SAQFVDASVGLAVGSMCLFIAGFAVGWGPVPLWLMSEIFPLHVKGATGICVLTNWL 419

DB 360 AAEFVDDVQVGLAVGSMCLFIAGFAVGWGPVPLWLMSEIFPLHVKGATGICVLTNWF 419

QY 420 MAFIVTKFSSLMELVRPYGAFWLASAFICFVSFLTLFCVPEPKGKTLQITAHFEGR 477

DB 420 MAFIVTKFNSVMEILRPYGAFWLTAAFCLSVLTLTVPETKGRTEQVTAHFEGR 477

RESULT 7

ABB89717
 ID ABB89717 standard; protein; 326 AA.

XX ABB89717;

DT 24-MAY-2002 (first entry)

XX Human polypeptide SEQ ID NO 2093.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 DE antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.

XX Homo sapiens.

XX WO200190304-A2.

XX 29-NOV-2001.

XX 18-MAY-2001; 2001WO-US016450.

XX 19-MAY-2000; 2000US-0205515P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;

XX WPI; 2002-122018/16.
 DR N-PSDB; ABL90126.
 XX
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 disorders.
 XX
 PS Claim 11; SEQ ID NO 2093; 2081pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB9040-ABB9044) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 326 AA;

Query Match 60.7%; Score 1491; DB 5; Length 326;
 Best Local Similarity 96.3%; Pred. No. 1.9e-134;
 Matches 289; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
 QY 164 MVVGIILAYLAGVLEWRWLVLCVPPSLMLLMCMFPTPFLTHQRRQEAALR 223
 Db 1 MVVGIILAYLAGVLEWRWLVLCVPPSLMLLMCMFPTPFLTHQRRQEAALR 60
 QY 224 FLWGEQWEDPPITGAEGSHLALLRQPGIYKPTIIGVSLMAFQOLSGVNAVMFAETIF 283
 Db 61 FLWGEQWEDPPITGAEGSHLALLRQPGIYKPTIIGVSLMAFQOLSGVNAVMFAETIF 120
 QY 284 BEAKFKDSSLASVVGVVQVLFVAALIMDRGRRLLLVSGVVMFSTSAFGAYFKLT 343
 Db 121 BEAKFKDSSLASVVGVVQVLFVAALIMDRGRRLLLVSGVVMFSTSAFGAYFKLT 180
 QY 344 QGGPNSHVAISAPVSAQPVDSVGLAWLAVGSMCLPIAGFVAGWGPIMPWMSEIFPL 403
 Db 181 QGGPNSHVAISAPVSAQPVDSVGLAWLAVGSMCLPIAGFVAGWGPIMPWMSEIFPL 240
 QY 404 HVKGVATGICVLTNWLMAFLVTKFSSLMELVRPYGAPWLASAFICFSVLTFLCVPETK 463
 Db 241 HVKGVATGICVLTNWLMAFLVTKFSSLMELVRPYGAPWLASAFICFSVLTFLCVPETK 300

RESULT 8
 AA06579
 ID AA06579 standard; protein; 262 AA.
 XX
 AC AA06579;
 XX

25-SEP-2001 (first entry)

Human protein having hydrophobic domain, HP10784.

Human; hydrophobic domain; gene therapy; nutritional supplement;
 cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;
 multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;
 haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;
 Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;
 haemostatic; thrombolytic; tumour growth inhibitor; anabolic;
 contraceptive; antinfertility; antinflammatory.

XX Homo sapiens.
 OS WO200149728-A2.
 XX
 PN 12-JUL-2001.
 PD
 XX
 PF 28-DEC-2000; 2000WO-JP009359.
 XX
 PR 06-JAN-2000; 2000JP-00000585.
 PR 06-JAN-2000; 2000JP-00000588.
 PR 11-JAN-2000; 2000JP-00002299.
 PR 03-FEB-2000; 2000JP-00026862.
 PR 03-MAR-2000; 2000JP-00058367.
 XX
 PA (PROT-) PROTEGENE INC.
 PA (SAGA) SAGAMI CHEM RES CENT.
 XX
 PI Kato S, Kimura T;
 XX
 DR WPI; 2001-418355/44.
 DR N-PSDB; AAD12574.
 XX
 PT Human proteins with hydrophobic domains and the nucleic acids encoding
 PT them, useful for preventing diagnosing and treating e.g. cancer,
 PT Alzheimer's and inflammation.
 XX
 PS Claim 1; Page 75; 563pp; English.
 XX
 CC The present sequence is human protein with hydrophobic domain, HP10784.
 CC The polynucleotide and polypeptide of the invention may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate polypeptide expression. The polynucleotides may be used to
 CC produce the polypeptide, by inserting the nucleic acids into a host cell
 CC and culturing the cell to express the protein. The polynucleotides and
 CC its complementary sequences may also be used as DNA probes in diagnostic
 CC assays and also used in gene therapy. The polypeptides may also be used
 CC as antigens in the production of antibodies and in assays to identify
 CC modulators of polypeptide expression and activity. The polypeptides and
 CC nucleic acids may be used as nutritional supplements, to modulate
 CC cytokine and cell proliferation activity, to modulate immune stimulation
 CC or suppression (e.g. for the treatment of microbial infections and
 CC autoimmune disorders such as multiple sclerosis, rheumatoid arthritis and
 CC insulin-dependent diabetes), to modulate haematopoiesis, to modulate
 CC tissue growth activity (e.g. for the treatment of Parkinson's disease,
 CC Huntington's disease and Alzheimer's disease), to modulate activin and
 CC inhibit activity (e.g. for controlling fertility), to modulate
 CC chemotactic and chemokinetic activity, to modulate haemostatic and
 CC thrombolytic activity, to modulate receptor ligand activity, to modulate
 CC inflammation and to inhibit tumour growth
 XX
 SQ Sequence 262 AA;

Query Match 46.2%; Score 1135; DB 4; Length 262;
 Best Local Similarity 91.4%; Pred. No. 2.2e-100;
 Matches 223; Conservative 2; Mismatches 7; Indels 12; Gaps 1;

QY 1 MTPEDPETOPLLLPGGSGAPRRVFLAFAAALGPLSGFALGYSPPAIPSLQRAAPP 60
 Db 1 MTPEDPETOPLLLPGGSGAPRRVFLAFAAALGPLSGFALGYSPPAIPSLQRAAPP 60
 QY 61 APRLDDAAASWFGAVVTLGAAAGVLCGLVDRAGRLKLLLCVSPVAVGFAVITAAQDV 120
 Db 61 APRLDDAAASWFGAVVTLGAAAGVLCGLVDRAGRLKLLLCVSPVAVGFAVITAAQDV 120
 QY 121 WMLLGGRLTLGLACGVASIVAPYVISEIAPVAVRGLLGSCVQLMVMVVGILLAYLAGWVLE 180
 Db 121 WMLLGGRLTLGLACGVASIVAPYVISEIAPVAVRGLLGSCVQLMVMVVGILLAYLAGWVLE 180
 QY 181 WRMLAVLCVPPSPSLMLLMCMFPTPFLTHQRRQEAALR 228
 Db 181 WRMLAVLCVPPSPSLMLLMCMFPTPFLTHQRRQEAALR 240

```
QY 229 EQGW 232
Db 241 DPGW 244

RESULT 9
ABU11283
ID ABU11283 standard; protein; 248 AA.
XX AC ABU11283;
XX XX
XX XX
Dt 10-FEB-2003 (first entry)
XX DE cDNA encoding human cancer suppressing protein P7425.
XX XX
XX KW Human; cancer suppressing protein; cancer.
XX XX
XX OS Homo sapiens.
XX XX
XX PN CN1351081-A.
XX PD 29-MAY-2002.
XX PF 31-OCT-2000; 2000CN-00127102.
XX PR 31-OCT-2000; 2000CN-00127102.
XX XX (SHAN-) SHANGHAI INST ONCOLOGY.
XX PA
XX PI Gu J;
XX XX
DR WPI; 2002-609437/66.
DR N-PSDB; ABX34029.
XX XX
PT New human protein with cancer cell growth suppressing function and a
PT polynucleotide encoding it, for treating diseases, such as, cancer.
XX XX
PS Claim 1; Page 23 (disclosure); 39pp; Chinese.
XX XX
CC This invention relates to the cDNA and protein sequences of a novel human
CC protein with cancer suppressing function. The invention also comprises a
CC method for preparing the polypeptide by recombination, and an application
CC of the polypeptide in treating diseases such as cancer, etc. Also
CC disclosed in an antagonist of the polypeptide and its medical action. The
CC present sequence represents a cancer suppressing protein of the invention
XX XX
SQ Sequence 248 AA;

Query Match 46.0%; Score 1131; DB 5; Length 248;
Best Local Similarity 93.0%; Pred. No. 5e-100;
Matches 227; Conservative 5; Mismatches 6; Indels 6; Gaps 2;

QY 164 MVVVGILLAYLAGVLEWRNLAVLGCVPVPSLMLLLCMFMPETPRFLLTQHRQEMAAALR 223
Db 1 MVVVGILLAYLAGVLEWRNLAVLGCVPVPSLMLLLCMFMPETPRFLLTQHRQEMAAALR 60

QY 224 FLWGSEQGWEDPPTGAEQSFHALLRQPGIYKPFIIIGVSLMAFQOLSGVNAVMYAETIF 283
Db 61 FLWGSEQGWEDPPTGAEQSFHALLRQPGIYKPFIIIGVSLMAFQOLSGVNAVMYAETIF 120

QY 284 EEAFKQSSLASVVGVIQVLTAAALIMDRAGRRLLLVLSGVVMVFSTSAFGAYFKLT 343
Db 121 EEAFKQSSLASVVGVIQVLTAAALIMDRAGRRLLLVLSGVVMVFSTSAFGAYFKLT 180

QY 344 QGGPENGSHVAISAPVSAQPDVDSVGLAWLAVGSMCLFIAGFVGWGP-IPWLLMSRIFP 402
Db 181 QGGPENGSHVAISAPVSAQPDVDSVGLAWLAVGSMCLFIAG- - - - -GPQALWSLLACLRF 235

QY 403 LHVX 406
Db 236 LHLQ 239

RESULT 10
AAB66938
ID AAB66938 standard; protein; 503 AA.
XX AC AAB66938;
XX XX
Dt 17-APR-2001 (first entry)
XX DE Rat GLUTX3.
XX XX
XX KW Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;
XX hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
XX hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX XX
XX OS Rattus sp.
XX XX
XX PN WO200104145-A2.
XX PD 18-JAN-2001.
XX PF 14-JUL-2000; 2000WO-IB001042.
XX PR 14-JUL-1999; 99US-0143907P.
XX PR 27-AUG-1999; 99US-0151140P.
XX PR 23-FEB-2000; 2000US-0184285P.
XX PR 13-JUL-2000; 2000US-00616132.
XX XX (UYLA-) UNIV LAUSANNE.
XX XX
XX PI Thorens B, Ibberson M, Uldry M;
XX XX
DR WPI; 2001-112615/12.
DR N-PSDB; AAF55871.
XX XX
PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
PT prevention, diagnosis and treatment of hexose transport disorders, e.g.
PT ischemia and diabetes.
XX XX
PS Claim 11; Page 82-83; 124pp; English.
XX XX
CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
CC AAB66932-AA66941). The GLUTX proteins are related to the facultative
CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC function. The GLUTX proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as ischaemia, diabetes,
CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is rat GLUTX3
XX XX
SQ Sequence 503 AA;

Query Match 39.5%; Score 969.5; DB 4; Length 503;
Best Local Similarity 43.6%; Pred. No. 4.2e-84;
Matches 212; Conservative 80; Mismatches 175; Indels 19; Gaps 7;

QY 3 PEDEETQPLLPGPGSAPGRVFLAAFAAALGPLSGFALGYSSPAISLQRAAPPAP 62
Db 20 PASPEE-----KARAGALQNRVFLATFAAVLGNFSGYALVVTSPVPAKKRSSDPAL 73

QY 63 RLDDAAASWFGAVVTLGAAAGGVLGWLVDRAGKLSLLCSVPFVAFVITAQDVVM 122
Db 74 RLDKIQASWFGSVFTLGAAAGLSLMLNDLLGRKLSIMFSAPVSAIGYALMAGARGLWM 133

QY 123 LLGRLTLTGACGVASLVAVPYISEIAYPAVRGLGSCVQLMVVVGILLAYLAGVLEWR 182
Db 134 LLLGRMLTGPAGLTACIPVYVSEIAPPVGRGALGATPQLMAVFGSLSLYALGLLLPWR 193

QY 183 WLAVLGCVPVSLMLLLCMFMPETPRFLLTQHRQEMAAALRFLW--GSEQGWEDPFI- 238
Db 194 WLAVAGEGPVLMILLFSFMNPGFRFLLSRDEALQAL--IWLRADSEVHWEFEQIQD 251

QY 239 --AEQSFHALLR--QPGIYKPFIIIGVSLMAFQOLSGVNAVMYAETIFFEAK-FKQSSL 293
Db 252 NVRRQSSRVSWAEAWEPVVRPILITVIMRFLOQLTGITPILVYLQITFSTSVLFSQQ 311
```


PT	ischemia and diabetes.	PF	07-JUL-2000; 2000EP-00114089.
XX		XX	
PS	Claim 11; Page 81-82; 124pp; English.	PR	08-JUL-1999; 99JP-00194486.
XX		PR	11-JAN-2000; 2000JP-00118774.
CC	The present invention relates to GLUTX proteins (AAF55865-AAF55871 and AAF66932-AAF66941). The GLUTX proteins are related to the facultative glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport function. The GLUTX proteins may be used in the diagnosis, prevention and treatment of hexose transport disorders such as ischaemia, diabetes, hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a neurodegenerative disease. The present sequence is human GLUTX3	XX	02-MAY-2000; 2000JP-00183765.
CC		XX	(HELI-) HELIX RES INST.
XX		XX	Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y, Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX		XX	WPI; 2001-524255/58.
XX		DR	N-PSDB; AAK94338.
SQ	Sequence 507 AA;	XX	
	Query Match 38.6%; Score 948; DB 4; Length 507;	PT	830 Primers useful for synthesizing full length cDNA clones and their use in genetic manipulation.
	Best Local Similarity 42.9%; Pred. No. 4.9e-82;	PS	Claim 8; SEQ ID NO 3034; 1380pp + Sequence Listing; English.
	Matches 217; Conservative 78; Mismatches 167; Indels 44; Gaps 11;	XX	
QY	10 QPLL-----PPGSAAPRGR-----RVFLAAFAALGPLSFGALGYSSPA 50	CC	The invention relates to primers for synthesising full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have been determined. Primers for synthesising the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a polypeptide encoded by a full length human cDNA of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from EPO
Db	3 EPLLGAEGPDYDTPPEKPPSPGDRARVGTQNKRVFLATFAVLGNFSGYALVITSPV 62	CC	
QY	51 IPSLQRAAPPAPRLDDAAASWFGAVVTLGAAAGVGLGGWLVDRAGRKLSLLCSVPFVAG 110	CC	
Db	63 IPALERSLDPDLHLTKSQASWFGSVFTLGAAGLSAMILNDLLGRKLSIMFSAVPSAAG 122	CC	
QY	111 FAVITAAQDVWMLLGGRLTLGLACGVASLVAPVYISIAIYPAVRGLLGSQVLMVVVGIL 170	CC	
Db	123 YAVMAGAHGLWMLLGRTLTGAGGLTAACIPVYVSEIAPPVGRGALGATPQLMAVFGSL 182	CC	
QY	171 LAYLAGWLEWRMLAVLGCVPSPSLMILLMCMPTPRFLTTOHRRQEAALRFLWSEQ 230	CC	
Db	183 SLVALGULLPWRMLAVAGAPVLIMILLSFMPNSPRFLLSRGRDEBALRALALWLRGIDV 242	CC	
QY	231 G--WEDPPIG---AEQSFHL--ALLRQPGIYKPIIGVSLMAFQOLSGVNAVMEFYAETIF 283	CC	
Db	243 DVHWEFGIQDNVRRQSSRVSWAEARAPHVCRPITVALLMELQLQLTGIPILVYLSIF 302	CC	
QY	284 EEAKF-----KDSSLASVVVGVIQVLTFAVALIMDRAGRLLLVLSGVWVFTSFAFGY 339	CC	
Db	303 DSTAVLLPPKDD---AAIVGXVRLLSVLIAALTMDLAGRKVLLFVSAAIMFAANLTGLY 359	CC	
QY	340 FKLTQGGPGNSSHVAISAPVS-----AQPVDASVG-LAMLVGSMCLFTAGFAVWGPI 392	CC	
Db	360 IHF---GPRPLSPNSTAGLESWGDLAQPLAAPAGYLTIVPLLATLWLFINGYAVWGPI 416	CC	
QY	393 PWLLMSEIFPLHVKGATGICVLTNWLMAFLVTKEFSSLMELVRPYGAFMLASAFICFSV 452	CC	
Db	417 TWLLMSEVPLRARGVASGLCVLASWLTAFVLTKSFLPVVSTFGLQVPFLFPAALCLVSL 476	CC	
QY	453 LFTLFCVPETKGTLEQITAHFE-GR 477	CC	
Db	477 VFTGCCVPETKGRSLEQIESFFRTGR 502	CC	
	RESULT 13		
	AAM93417		
ID	AAM93417 standard; protein; 507 AA.		
XX			
AC			
XX			
DT	06-NOV-2001 (first entry)		
XX			
DE	Human polypeptide, SEQ ID NO: 3034.		
XX			
KW	Human; full length cDNA; cDNA synthesis; oligo-capping.		
XX			
OS	Homo sapiens.		
XX			
FN	EP1130094-A2.		
XX			
PD	05-SEP-2001.		
XX			

```

RESULT 14
ADA84077
ID ADA84077 standard; protein; 507 AA.
XX AC ADA84077;
XX DT 20-NOV-2003 (first entry)
XX DE Human SLC2A6 protein.
XX KW human; marker; expressed sequence tag; EST; arabidopsis; tumour;
XX KW stress-induced phenotype; hyperosmotic stress; colon cancer; immunogen;
XX OS Homo sapiens.
XX PN WO2002103028-A2.
XX PD 27-DEC-2002.
XX PF 30-MAY-2002; 2002WO-IB004189.
XX PR 30-MAY-2001; 2001US-0293999P.
XX PR 22-OCT-2001; 2001US-0330457P.
XX PR 19-FEB-2002; 2002US-0357144P.
XX PA (BIOM-) BIOMEDICAL CENT.
XX PI Baranova AV, Yankovsky NK, Kozlov AP, Lobashev AV, Krukovskaya LL;
XX DR WPI; 2003-175241/17.
XX DR N-PSDB; ADA84076.
XX PT Determining if a nucleic acid is a marker for a phenotype/cell type of
XX PT interest, by global comparison of expressed sequence tags known to be
XX PT expressed in the phenotype/cell type with all ESTs expressed in normal
XX PS Claim 29; Page 458-460; 516pp; English.
XX CC The invention relates to a novel method for determining if a nucleic acid
XX CC is a marker for a predetermined phenotype/cell type of interest from a
XX CC biological species. The method comprises performing a global comparison
XX CC of a group of expressed sequence tags (ESTs) known to be expressed in the
XX CC phenotype/cell type of interest with all ESTs expressed in normal tissue
XX CC in order to identify ESTs that are preferentially expressed in the
XX CC phenotype/cell of interest. A method of the invention is useful for
XX CC determining whether a nucleic acid is a marker for a predetermined
XX CC phenotype or cell type of interest from a biological species, preferably
XX CC Arabidopsis or human. The cell type of interest is an abnormal cell such
XX CC as a tumour cell, and the predetermined phenotype is a stress-induced
XX CC phenotype such as hyperosmotic stress or high salt conditions. A method
XX CC of the invention is also useful for determining the progression of colon
XX CC cancer in a human, for detecting a tumour cell, and for regulating or
XX CC preventing the growth of a tumour cell. An antibody of the invention is
XX CC useful for detecting the absence or presence of peptides encoded by
XX CC tumour-associated markers. A polypeptide of the invention is useful as an
XX CC immunogen for vaccinating an animal. The present sequence represents a
XX CC tumour-associated antigen of the invention.
XX SQ Sequence 507 AA;

Query Match 38.6%; Score 948; DB 6; Length 507;
Best Local Similarity 42.7%; Pred. No. 4.9e-82;
Matches 216; Conservative 80; Mismatches 166; Indels 44; Gaps 11;

Qy 10 QPLLG-----PPGGSAPRGR-----RVFLAFAAALGPLSFALGYSSPA 50
Dy 3 EPLLGAAGPDYDTFPPKPPSPGDRARVGTLNQKRVFLATFAVLGNFSGYALVYTSFV 62
Qy 51 IPSLQRAAPNAPRLDDAASWFGAVVTLLGAAGVGLGWLVDRAGRKLILLCSVPFVAG 110

```

```

Db 63 IPALERSLDPDHLHTKSQASWFGSVFTLLGAAGGLSAMILLDLGRKUSIMFSAPVPSAAG 122
Qy 111 FAVITAAQDVWMLLGGRLTLGLACVASLAVPVIYSEIAYPAVRGLGLGSCVOLMVVVGIL 170
Db 123 YALMAGAHGLWMLLGLRTLTGFAGGLTAACIPVYVSEIAPPVGRGALCATPOLMAVFGSL 182
Qy 171 LAYLAGWLEWRWLAIVLCVPPSLMLLLCMPEPTEPRELLTQHRQEAAMALRFLWGSEQ 230
Db 183 SLVALGGLLPWRWLAIVAGEAEVLIMILLSPMSPREFLLSRGRDEEALRALAWRGTDV 242
Qy 231 G--WEDPPIG--AEQSPHL--ALLRQPGIVKPEFTIGVSLMAFOQLSGVNAVMEVAETIP 283
Db 243 DVHFEFEIQNVRRQSRVSWAERAPHVCRTITVALMLRLQLQLTGITTLVTLQISIF 302
Qy 284 BEAKF-----KDSLSASVVVGVQVLTAVAAALIMDRAGRLLLLVLSGVVMVFSTSAFGAY 339
Db 303 DSTAVLLPPKDD--AAIVGAVRLLSVLIAALTMDLAGRKVLLFVSAAIMFAANLTGLY 359
Qy 340 FKLTOGGPGNSSHVAISAPVS-----AOPVDASVG--LAWLAVGSMCLFIAGFAVWGPI 392
Db 360 IHF---GPRPLSPNSTAGLSESGDIAQPLAAPAGYTLTVPLLATMLFIMGYAVWGPI 416
Qy 393 FWLLMSEIFPLHVKGVATGICVLTNWLMAFVLTKFEFSSIMEVLRPYGAFWLASAFICFSV 452
Db 417 TWLLMSEVLPARGVASGLCVLASWLTAFVLTKSFLPVVSTFGLQVDFPFPAALICIVSL 476
Qy 453 LFTLFQVPEPKGKTLQITAHFE-GR 477
Db 477 VFTGCCVPETKGRSLEQIESPFTRG 502

RESULT 15
ABP58365
ID ABP58365 standard; protein; 507 AA.
XX AC ABP58365;
XX DT 07-APR-2003 (first entry)
XX DE Human solute carrier type 2A polypeptide 8923733.
XX KW Human; solute carrier type 2A; SLC2A; glucose transporter; p53; cancer;
XX KW cytostatic; gene therapy.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Domain 39..61 /note= "transmembrane domain"
FT Domain 40..498 /note= "sugar transporter domain"
FT Domain 81..103 /note= "transmembrane domain"
FT Domain 110..132 /note= "transmembrane domain"
FT Domain 136..158 /note= "transmembrane domain"
FT Domain 165..187 /note= "transmembrane domain"
FT Domain 191..213 /note= "transmembrane domain"
FT Domain 279..301 /note= "transmembrane domain"
FT Domain 316..333 /note= "transmembrane domain"
FT Domain 340..362 /note= "transmembrane domain"
FT Domain 391..413 /note= "transmembrane domain"
FT Domain 426..448 /note= "transmembrane domain"
FT Domain 458..480 /note= "transmembrane domain"

```


Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 19:22:56 ; Search time 45 Seconds

(without alignments)
1019.630 Million cell updates/sec

Title: US-09-886-954A-1

Perfect score: 2457

Sequence: 1 MTPEDPEETQPLIGPPGSA.....CVPETKGKTLQITAHFEGR 477

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	734	29.9	487	2 E96782	hypothetical prote
2	722	29.4	490	2 T14545	probable sugar tra
3	607	24.7	461	2 D70073	metabolite transpo
4	607	24.7	471	2 A80868	L-arabinose isomer
5	606.5	24.7	472	2 S47089	arabinose-proton s
6	603	24.5	472	2 B26430	L-arabinose isomer
7	603	24.5	472	2 B91091	L-arabinose isomer
8	603	24.5	472	2 E85936	L-arabinose isomer
9	591.5	24.1	464	2 F65079	galactose-proton s
10	587.5	23.9	464	2 C91106	galactose-proton s
11	587.5	23.9	464	2 F85951	galactose-proton s
12	584.5	23.8	496	2 T52132	probable sugar tra
13	584.5	23.8	575	2 T43400	myo-inositol trans
14	583.5	23.7	464	2 A80877	galactose-proton s
15	582	23.7	521	2 G84864	probable membrane
16	574.5	23.4	493	2 S38981	glucose transport
17	570.5	23.2	496	2 A31986	glucose transport
18	569.5	23.2	493	2 A41751	glucose transport
19	567	23.1	580	2 B86426	hypothetical prote
20	560.5	22.8	457	2 E70070	metabolite transpo
21	556.5	22.6	560	2 T51485	sugar transporter-
22	554.5	22.6	464	2 F69587	L-arabinose transp-
23	551	22.4	461	2 G85059	probable sugar tra
24	548.5	22.3	523	2 S25015	monosaccharide tra
25	546.5	22.2	526	2 T01853	probable hexose tr
26	546	22.2	491	2 A26430	xylose transport p
27	546	22.2	491	2 F91255	xylose-proton symp
28	546	22.2	491	2 B86096	xylose-proton symp
29	542.5	22.1	522	2 A31556	glucose transport

ALIGNMENTS

RESULT 1

E96782

hypothetical protein F22H5.6 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana [mouse-ear cress]

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C:Accession: E96782

R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E96782

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-487 <STO>

A:Cross-references: GB:AB005173; NID:g10092276; PIDN:AAG12689.1; GSPDB:GN00141

C:Genetics:

A:Gene: F22H5.6

A:Map position: 1

C:Superfamily: glucose transport protein

Query Match 29.9% Score 734; DB 2; Length 487;

Best Local Similarity 37.5%; Pred. No. 8.9e-47;

Matches 173; Conservative 77; Mismatches 177; Indels 34; Gaps 6;

Qy	26	VFLAAFAAALGPLSGFALGYSSPAIFSLQRAAPPAPRLDDAAASWFGAVVTLGAAGGV	85
Db	48	VLACVLIVAGPIQFGTCGVSSPT---QAAIKDLGLTVSEYSVFGLSNVGMVGAI	103
Qy	86	LGGWLVDPAGKLSLLCSVPFVAGFVITAAQDVMWLLGGRILTLGLACVASLVAPYVI	145
Db	104	ASGQIAEYIGRKGSLMTAAIPNIIGLCISPAKDTSLYMGRLLEGFVGFIISYTVPYVI	163
Qy	146	SEIAYPAVRGLGSCVOLVMVVGILLAYLAGWLEWRLAVLGCVPSPSLMLLMCFMPET	205
Db	164	AEIAPQNRGGLSGVNLQSVTIGIMLAYLLGLFVPRILAVLGLPCTLLTPGFFIPES	223
Qy	206	PRFLLTQHRROEAMAAFLFLWGSEQGMEDDPDPAEQ-----SFHALLRQPGI	253
Db	224	PRWLAKGMTDTEFETSLQVLRGFE--TDITVEVNEIKRSVASTKRTNTVFDLKRKY	280
Qy	254	YKPIIIGVSLMAFOQLSGVNAWMPYAEITPEAKFKSSLASVVGVIQVLFVAALIM	313
Db	281	YFPLMVIGILLVLOQLGGINGVLFYSSTFSSAGVTSNAATFGVGAIQVVATAISTWIV	340
Qy	314	DRAGRRULLVLSGVVMVFSTSAFGAYFKLTGGFGCNSHVAISAPVSAQPDVASVGLAWL	373

Db 341 DRAGRRLLITSSVGMTISLVIAAAYLKE-----FVSPDS-----DMYSWLSIL 386
QY 374 AVGSMCLFTAGFVAGVGPWPWLLMSEIPLHVKGVATGICVLTNNMLAFVTKYFSSLME 433
Db 387 SVVGVVAVVFFSLGMPWLMIMSEIPLVNIKGLAGSIATLANFFSWLIITMT-ANLL 445
QY 434 VLRPYGAFLASAFICFISVLTFCVPTKGTLEQITAHF 474
Db 446 AWSGGTTLTGLVCAFTVVFVTLVWPETKGTLELQSLF 486

RESULT 2
T14545
Probable sugar transporter protein - beet
C:Species: Beta vulgaris (beet)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T14545
R:Chou, T.J.; Bush, D.R.
Plant Physiol. 110, 511-520, 1996
A:Title: Molecular cloning, immunochemical localization to the vacuole, and expression
A:Reference number: Z18131; MUID:96351183; PMID:8742332
A:Accession: T14545
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-490 <CHI>
A:Cross-references: EMBL:U43629; NID:gl209755; PIDN:AAB53155.1; PID:gl209756
A:Experimental source: tonoplast
C:Superfamily: glucose transport protein
C:Keywords: transmembrane protein

Query Match 29.4%; Score 722; DB 2; Length 490;
Best Local Similarity 36.1%; Pred. No. 6.9e-46;
Matches 167; Conservative 81; Mismatches 177; Indels 38; Gaps 5;
QY 26 VFLAFAAALGPLSGFALGYSSPAIPSLQRAAPPAPRLDDAAASWFGAVVTLGAAAGV 85
Db 50 VLACVLI VALGPIQFGFTAGYSPT-----QSAITNELGLSVAEYSWFGSLSNVGAMVCAI 105
QY 86 LGGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDVWMLLGRLLTGLACGVASLVAPVYI 145
Db 106 ASGQISEYIGKKGSLMIAIENIIGWLAI SFAKDSFLYMGRLGEGVGVIISVTPVYI 165
QY 146 SEIAPAVRGLLGSCVOLMVVGVGILLAYLAGVLEWRWLVLCVPPSLMLLMCFWPET 205
Db 166 SEIAPQNLRGALGSVNQSLVIGIMLSYMLGLFVPWRILAVLGPCTILIPGFFIPES 225
QY 206 PRFLTQHRROEAMALRFLMGSGQGWEDDPIGAE-----QSFDLALLRQP 251
Db 226 PRWLAKGMMEPEFVSLQVLRGF-----DTDISLEVNEIKRSVASSSKRTTIRFAELRQ 280
QY 252 GIYKPFITIGVSLMAFQOLSGVNAVMEYAEITFPEAKFKDSSLASVVGVVTOVLETAAL 311
Db 281 RYWLPMIENGLLTIQQUSGINGVLFYSSTIFKEAGVTSSNAATFGLGAVQVIAVTVTW 340
QY 312 IMDRAGRRLLLVLSGVVMVFTSFAFGAYFKLTQGGPGNSHVAISAPVSAQPDASVGLA 371
Db 341 LVDSKSGRRLLIVSSGWTLSLLVAVMSFFLKENVSDSEIWSV-----FS 386
QY 372 WLAVGSMCLFTAGFVAGVGPWPWLLMSEIPLHVKGVATGICVLTNNMLAFVTKYFSS 431
Db 387 ILSVGVVAVVTFSLGIGAIPTWIMSEIPLINIKGLAGSIATLANFVAVIWTMT-ANI 445
QY 432 MEVLRPYGAFLASAFICFISVLTFCVPTKGTLEQITAHF 474
Db 446 MLSWNSGTFISIYVAVVCAFTVVFVWIWPETKGTLEIQNSF 488

RESULT 3
D70073
metabolite transport protein homolog yxcC - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C:Accession: D70073
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berke
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Frit, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Ser
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Yata, K.; Yoshi
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yotsu, V.; Uchiyama
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: D70073
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-461 <KUN>
A:Cross-references: GB:Z99124; GB:AL009126; NID:g2636442; PIDN:CAB16017.1; PID:g2636527
A:Experimental source: strain 168
C:Genetics:
A:Gene: yxcC
C:Superfamily: glucose transport protein

Query Match 24.7%; Score 607; DB 2; Length 461;
Best Local Similarity 31.1%; Pred. No. 2e-37;
Matches 146; Conservative 85; Mismatches 192; Indels 46; Gaps 7;
QY 25 RVFLAFAAALGPLSGFALGYSSPAIPSLQRAAPPAPRLDDAAASWFGAVVTLGAAAGG 84
Db 6 RKMIMYFPGALGGLLYGDTGVISGALLFINNDIPLTLTLEGLVVS-----MLLGAIFGS 61
QY 85 VLGGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDVWMLLGRLLTGLACGVASLVAPVY 144
Db 62 ALSTGCDRWRGRRKVVFLSIIIFIIGALACAFSQTIGMLIASRVILGAVGGSTALVDPVY 121
QY 145 ISEIAPAVRGLLGSCVOLMVVGVGILLAYLAGVWL-----BWRWLAVLCVPPSLMLLM 200
Db 122 LSEMAPTKIRGTGTNNLMIVTIGILLAYIVNLYFTPEAWRMVGLAAVPAVLLIGIA 181
QY 201 FMPETPRFLTQHRROEAMALRF-----LWSEQGWEDDPIGAEOSFHALLRQ 250
Db 182 FMPESPWLVRGSEEEARRIMNITHDPKDIEMELAEMKQGEAE-----KKTTLGLVKA 236
QY 251 PGYKPFITIGVSLMAFQOLSGVNAVMEYAEITFPEAKF--KDSLSASVVGVVTOVLETA 308
Db 237 KNIRPMLLIGVGLAIFQAVGINTVIYVPTIIFKAGLGTSASALGTNGIGILNVICIT 296
QY 309 AALIMDRAGRRLLLVLSGVVMVFTSFAFGAYFKLTQGGPGNSHVAISAPVSAQPDASV 368
Db 297 AMILIDVRGKRKLLIWSGVGITLSLAALSGVL-LTLG-----LSA 335
QY 369 GLAWLAVGSMCLFTAGFVAGVGPWPWLLMSEIPLHVKGVATGICVLTNNMLAFVTKY 428
Db 336 STAWMTVVFVGVVIFVYQATWGPVVMVLMPELFPKARGAATGFTTILVLSAANLIVSLV 395
QY 429 SSLMEVLRPYGAFLASAFICFISVLTFCVPTKGTLEQITAHFGR 477
Db 396 PLMSANGIAWVFWFVSVICLLSFFFAFVNVPTKGSLEIEASLKKR 444

RESULT 4
AB0868
L-arabinose isomerase [imported] - Salmonella enterica subsp. enterica serovar Typhi (str
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AB0868
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AB0868
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-471 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD02842.1; PID:g16504096; GSPDB:GN00176
C;Genetics:
A;Gene: STY3160
C;Superfamily: glucose transport protein

Query Match 24.7%; Score 607; DB 2; Length 471;
Best Local Similarity 32.7%; Pred. No. 2.1e-37;
Matches 158; Conservative 85; Mismatches 192; Indels 48; Gaps 10;
QY 1 MTPEDPEETQPLGPGGSAPGRRRVFLAAFAALGFLSPGALGYSSPAIPSLQRAAPP 60
Db 9 LTRSLURDT-----RRNMFFVSVAAG-LFLGLDIGIAGALPFIITDHFVL 55
QY 61 APRLDDAAASWFGAVVTGLGAAAGVLLGWLVDRAGRKLSLLLSGVFPFVAGFAVITAAQDV 120
Db 56 TSLQEQ-----WVYSSMMMLGAAIGALFNGWLSFRLGRKYSLMAGAILFVLGSLGSAFASV 111
QY 121 WMLLGRLLTGLACGVASLAVPYVISEIAYPAVRGLLGSCVQLMVPVGVGILLAYLAGWLE 180
Db 112 EVLIGARVILGAVGIASTYAPLYLSEMASENVRGKMISMYQLMVTILGIVLAFSLDTAFS 171
QY 181 W--RWLAVLG--CVPPSLMLLMCFEPETPRELLTQHRROEAMAAALFRLFGSQGWEDPP 236
Db 172 YSGNRAMUGVLAIPAVLLILVVELPNSRWLAQAQRHIEAEVRLMLRDTSEKARDEL 231
QY 237 IGAEQSFHL-----ALLR-QPGIYKPFIIIGVSLMAFQQLSGVNAVIFYAEIIFEEAKF-- 288
Db 232 NEIPRESLKLQGGNALFKANNRRAVFLGMLLQAMQQTGMNIIIMYAPRIFKMGAGTT 291
QY 289 -KSSLASVVGVIQVLFVAALIMDRAGRLLLVLGVVNVFVSFAFGAYFKLTQGGP 347
Db 292 TEQQTATLVVGLTFEATPIAVTVDKAGKPKALKIGFSVMALGTLVLG--YCLMQFDN 349
QY 348 GNSHVAISAPVSAQPDVDSVGLAWLAVGSMCLFIAGFVAGWGPIPWLLMSIEIFELHVK 407
Db 350 GT-----ASSGLSWLSVGTMMCIAGYAMSAAPVVMWILCSEIQPLKCRD 393
QY 408 VATGICVLTNWLMFLVTKFSSLMELVRPYGAFWLASAFICISVLTFLFCVPETKGTLL 467
Db 394 FGITCSTTTNWSNMIIAGATFLTLDSIGAAGTFWLYTALNIAFIGITFWLIPETKNVTL 453
QY 468 EQI 470
Db 454 EHI 456

RESULT 5
S47089
arabinose-proton symporter - Klebsiella oxytoca
C;Species: Klebsiella oxytoca
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C;Accession: S47089
R;Shatwell, K.P.; Charalambous, B.M.; McDonald, T.P.; Henderson, P.J.F.
submitted to the EMBL Data Library, June 1994
A;Description: The nucleotide sequence of the gene araE for arabinose-proton symport in
A;Reference number: S47089
A;Accession: S47089
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-472 <SHA>
A;Cross-references: EMBL:X78598; NID:9498919; PIDN:CAA56110.1; PID:g498920
C;Superfamily: glucose transport protein
C;Keywords: arabinose transport

Query Match 24.7%; Score 606.5; DB 2; Length 472;
Best Local Similarity 33.7%; Pred. No. 2.3e-37;
Matches 153; Conservative 84; Mismatches 182; Indels 35; Gaps 8;
QY 30 AFPAALGPLSPFGALGYSSPAIPSLQRAAPPAPRLDDAAASWFGAVVTGAAAGVGLGW 89
Db 26 SIAAAVAGLLFGLDIGIAGALPFIITDHFVLSRLQE-----WVYSSMMMLGAAIGALFNGW 81
QY 90 IYDRAGRKLSLLCSVPFVAGFAVITAAQDVMMLLGRLLTGLACGVASIVAPVYISETA 149
Db 82 LSFRLGRKYSMLVMGAVLFVAGSVGSFAFATSEVEMLLVARIVLGVAVGIASTYAPLYSEMA 141
QY 150 YPAVRGLLGSCVQLMVPVGVGILLAYLAGWLEW--RWLAVLG--CVPPSLMLLMCFMPET 205
Db 142 SENVRGKMISMYQLMVTILGIVMAFLSDTAFSYSGNWRMLGVLAIPAVVLIILVIFLPS 201
QY 206 PRELLTQHRROEAMAAALFRLFGSQGWEDPPPIGAEOSFHL-----ALLR-QPGIYKPFII 259
Db 202 PRLAEBKGRHVEAEVRLMLRDTSEKARDELNEIRESLKLQGGWALFKVNRNVRRAVFL 261
QY 260 GYSLMAFQQLSGVNAVIFYAEIIFEEAKF---KSSLASVVGVIQVLFVAALIMDRA 316
Db 262 GMLLQAMQQTGMNIIIMYAPRIFKMGAGTTTEQQMVATLVVGLTFEATPIAVTVDKA 321
QY 317 GRLLLVLSGVNVFVSFAFGAYFKLTQGGPGNSHVAISAPVSAQPDVDSVGLAWLAVG 376
Db 322 GRKPKALKIGFSVMAIGTLVLG--YCLMQFDNGT-----ASSGLSWLSVG 363
QY 377 SMCFLTAGFVAGWGPIPWLLMSIEIFELHVKVAGTCVLTNWLMAFLVTKFSSLMELVR 436
Db 364 MTMCCIAGYAMSAAPVVMWILCSEIQPLKCRDFIGTCTTTNWSNMIIAGATFLTLDAIG 423
QY 437 PYCAPWLASAFICISVLTFLFCVPETKGTLEOI 470
Db 424 AAGTFWLYTALNVAFIGVTFWLPETKNVTLLEHI 457

RESULT 6
B26430
L-arabinose isomerase (EC 5.3.1.4) - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 01-Mar-2002
C;Accession: B26430; A28075; I40996; B65067
R;Maiden, M.C.J.; Davis, E.O.; Baldwin, S.A.; Moore, D.C.M.; Henderson, P.J.F.
Nature 325, 641-643, 1987
A;Title: Mammalian and bacterial sugar transport proteins are homologous.
A;Reference number: A93389; MUID:87115869; PMID:3543693
A;Accession: B26430
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-472 <MAI>
R;Maiden, M.C.J.; Jones-Mortimer, M.C.; Henderson, P.J.F.
J. Biol. Chem. 263, 8003-8010, 1988
A;Title: The cloning, DNA sequence, and overexpression of the gene araE coding for arabi
A;Reference number: A28075; MUID:88228015; PMID:2836407
A;Accession: A28075
A;Molecule type: DNA
A;Residues: 1-472 <MA2>
A;Cross-references: GB:J03732; NID:g145320; PIDN:AAA23469.1; PID:g145321
R;Stoner, C.; Schleif, R.
J. Mol. Biol. 171, 369-381, 1983
A;Title: The araE low affinity L-arabinose transport promoter. Cloning, sequence, trans
A;Reference number: I40996; MUID:84114888; PMID:6319708
A;Accession: I40996
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-25,'Y',349,'R' <RES>
A;Cross-references: EMBL:X00272; NID:g40940; PIDN:CAA25075.1; PID:g40941
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: B65067
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-472 <BLAT>
 A;Cross-references: GB:AE000368; GB:U00096; NID:92367165; PID:AACT5980.1; PID:g1789207;
 A;Experimental source: strain K-12, substrain MG1655
 C;Genetics:

A;Gene: araE
 A;Map position: 61 min
 A;Superfamily: glucose transport protein
 C;Keywords: arabinose transport; intramolecular oxidoreductase; isomerase; membrane protein

Query Match 24.5%; Score 603; DB 2; Length 472;
 Best Local Similarity 33.3%; Pred. No. 4.1e-37;
 Matches 158; Conservative 84; Mismatches 197; Indels 36; Gaps 9;

```

QY 9 TQPLLGPAGRRVFLAFAAALGFLSFGFALGYSSPAISLQRAAPPAPRLDDAA 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 6 TESALTTPSLRDTTRMMNFVSV-AAAAGLLFGLDIGIAGALPFIIDHFLVTSRLQ-- 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 69 ASWFGAVVTGAAAAGVGLGWLVDRAGRKLSLLCSVPFVAGFAVITAQDVWMLLGGRL 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 --WVSSMMLGAAIGALFNGLSPRLGRKYSLMAGAILFVLGSGSAPATSVEMLIAARV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 129 LTGLACGVASLVAPVYISEIAYPAVRGLGSCVOLMVVGVGILLAYLAGWLEW--RWLAV 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 VLGIAGVGIASVTAPLYLSEMASENVKGMISMYQLMTVLGIVLAFSLDTAFSYSGNWRAM 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 LG--CVPSMLLMCMPEPTPRLLTQHRROEAMALRFLWGSEQWEDDPIGAEQSFH 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 LGVIALPAVLIIILVFLNPSRWLABKGRHIEAEVLRMLRDTSEKAREELNIRESLK 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 245 L-----ALLR-QPGIYKPFIIIGVSLMAFQQLSGVNAVVFYAEITFEBAKF---KDSSIAS 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 LKQGGWALFKINRVRRAVFLGMLLQAOQFTGMNIIIMYAPRIFKMGFTTTTQQMIAT 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 296 VVGVIOVLFTVAALIMDRAGRLLLVLSGVNVVFTSAPGAFYFKLTQGGPGNSSHVAI 355
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 LVWGLTFMEATFIAVFTVDKAGRPALKIGFSVVALGTLVLG--YCLMQPDNGT----- 352
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 356 SAPVSAQPVDAVGLAMVAGSMCLFTAGFAVGVGPPIWLLMSIFFLHVKGVATGICVL 415
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 353 -----ASSGLSWLSVGMTCIAGYMSAAPVWILCSRIQPKRCDFGTCSTT 402
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 416 TNMLMAFLVTKFESSLMEVLRPGAFWLASAFICFISVLFVFCVPETKGTLEOI 470
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 403 TNWVSNMIIGATFTLLDSIGAAGTFWLYTALNIAFVGITFWLIPETKNTVLEHI 457
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 7
 B91091
 L-arabinose isomerase (EC 5.3.1.4) - Escherichia coli (strain O157:H7, substrain RIMD 05
 C;Species: Escherichia coli
 C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 31-Dec-2001
 C;Accession: B91091
 R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
 A;Reference number: A99629; MUID:21156231; PMID:11258796
 A;Accession: B91091
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-472 <HAY>
 A;Cross-references: GB:BA000007; PID:BA037121.1; PID:g13363170; GSPDB:GN00154
 A;Experimental source: strain O157:H7, substrain RIMD 050952
 C;Genetics:

A;Gene: Ecs3698
 A;Map position: 61 min
 A;Superfamily: glucose transport protein
 C;Keywords: intramolecular oxidoreductase; isomerase

Query Match 24.5%; Score 603; DB 2; Length 472;
 Best Local Similarity 33.3%; Pred. No. 4.1e-37;

```

Matches 158; Conservative 84; Mismatches 197; Indels 36; Gaps 9;
QY 9 TQPLLGPAGRRVFLAFAAALGFLSFGFALGYSSPAISLQRAAPPAPRLDDAA 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 6 TESALTTPSLRDTTRMMNFVSV-AAAAGLLFGLDIGIAGALPFIIDHFLVTSRLQ-- 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 69 ASWFGAVVTGAAAAGVGLGWLVDRAGRKLSLLCSVPFVAGFAVITAQDVWMLLGGRL 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 --WVSSMMLGAAIGALFNGLSPRLGRKYSLMAGAILFVLGSGSAPATSVEMLIAARV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 129 LTGLACGVASLVAPVYISEIAYPAVRGLGSCVOLMVVGVGILLAYLAGWLEW--RWLAV 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 VLGIAGVGIASVTAPLYLSEMASENVKGMISMYQLMTVLGIVLAFSLDTAFSYSGNWRAM 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 LG--CVPSMLLMCMPEPTPRLLTQHRROEAMALRFLWGSEQWEDDPIGAEQSFH 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 LGVIALPAVLIIILVFLNPSRWLABKGRHIEAEVLRMLRDTSEKAREELNIRESLK 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 245 L-----ALLR-QPGIYKPFIIIGVSLMAFQQLSGVNAVVFYAEITFEBAKF---KDSSIAS 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 LKQGGWALFKINRVRRAVFLGMLLQAOQFTGMNIIIMYAPRIFKMGFTTTTQQMIAT 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 296 VVGVIOVLFTVAALIMDRAGRLLLVLSGVNVVFTSAPGAFYFKLTQGGPGNSSHVAI 355
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 LVWGLTFMEATFIAVFTVDKAGRPALKIGFSVVALGTLVLG--YCLMQPDNGT----- 352
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 356 SAPVSAQPVDAVGLAMVAGSMCLFTAGFAVGVGPPIWLLMSIFFLHVKGVATGICVL 415
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 353 -----ASSGLSWLSVGMTCIAGYMSAAPVWILCSRIQPKRCDFGTCSTT 402
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 416 TNMLMAFLVTKFESSLMEVLRPGAFWLASAFICFISVLFVFCVPETKGTLEOI 470
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 403 TNWVSNMIIGATFTLLDSIGAAGTFWLYTALNIAFVGITFWLIPETKNTVLEHI 457
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 8

B85936
 L-arabinose isomerase (EC 5.3.1.4) - Escherichia coli (strain O157:H7, substrain EDL933
 C;Species: Escherichia coli
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Dec-2001
 C;Accession: B85936
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca
 Nature 409, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A;Reference number: A85480; MUID:21074935; PMID:11206551
 A;Accession: B85936
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-472 <STO>
 A;Cross-references: GB:AE005174; NID:g12517333; PIDN:AGS7953.1; GSPDB:GN00145; UWGP:Z41
 A;Experimental source: strain O157:H7, substrain EDL933
 C;Genetics:

A;Gene: araE
 C;Superfamily: glucose transport protein
 C;Keywords: intramolecular oxidoreductase; isomerase
 Query Match 24.5%; Score 603; DB 2; Length 472;
 Best Local Similarity 33.3%; Pred. No. 4.1e-37;
 Matches 158; Conservative 84; Mismatches 197; Indels 36; Gaps 9;

```

QY 9 TQPLLGPAGRRVFLAFAAALGFLSFGFALGYSSPAISLQRAAPPAPRLDDAA 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 6 TESALTTPSLRDTTRMMNFVSV-AAAAGLLFGLDIGIAGALPFIIDHFLVTSRLQ-- 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 69 ASWFGAVVTGAAAAGVGLGWLVDRAGRKLSLLCSVPFVAGFAVITAQDVWMLLGGRL 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 --WVSSMMLGAAIGALFNGLSPRLGRKYSLMAGAILFVLGSGSAPATSVEMLIAARV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 129 LTGLACGVASLVAPVYISEIAYPAVRGLGSCVOLMVVGVGILLAYLAGWLEW--RWLAV 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 VLGIAGVGIASVTAPLYLSEMASENVKGMISMYQLMTVLGIVLAFSLDTAFSYSGNWRAM 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```


QY 125 GGRLLTGLACGASVAVPVVISEIAYPAVRGLGSCVQLMVMVVGILLAYLAGWVLE--- 180
Db 185 VGRFVIGWGVGIASLIPIYLSIETAPSKIRGRVIVVLLITAGQVITAGIDTAFERHVN 244
QY 181 -WRWLAVLGVCPVPSLMLLMCFMPEIPREFLLTQHRROEAMALRFLWGS----- 228
Db 245 GWRWVGLAMVPAFQFILLWLPSPRLVLKRSQEAVENTLARIYPTAHPEYIKTKLY 304
QY 229 --EQWEDDPIGAE-----QSFHALLRQPGIYKPFITIGVSLMAFQOLSGVNAVMAFET 281
Db 305 LIQEGVRDPFSGSRWQIVKTFK-ELYFNPDSFRALILACGLQAMQOLSGFNSLMVPSST 363
QY 282 IFBEAKFKDSSLASVVGVVQLVFTAVAAALIMDRAGRELLVLS--GVWVFTSAGAY 339
Db 364 IFEVGNNPTATGLIIIAANFVETIVAFGVDFGGRILLLLTVWGMIAALIVCAVAFH 423
QY 340 FKLTQGGPGNSHVAISAPVSAQPDASVGLAVLAVG--SMCLFIAGFVAGWGPPIWLLM 397
Db 424 F-LPKDENGNYT-----SGQSNAMAIIVLISMIYVVAASGLGNLPW-QQ 467
QY 398 SEIPLPHVKGVATGICVLTNWMAFLVTKBFSSLMVLRPYGAPWLASAFICFSLVETLP 457
Db 468 SELFPMSVRGLGCMSTAVNWAGNLGIGASFLTMSBITPTGTPALYGLGCLFLGWLGA 527
QY 458 CVPETKGTLEQI 470
Db 528 CYPDLTDYTIIEI 540

RESULT 14
AC0877
Galactose-proton symport (galactose transporter) STY3244 [imported] - Salmonella enterid
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AC0877
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Ouail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AC0877
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-464 (PAR)
A;Cross-references: GB:AL513382; PIDN:CAD02915.1; PID:g16504168; GSPDB:GN00176
C;Superfamily: Glucose transport protein

Query Match 23.7%; Score 583.5; DB 2; Length 464;
Best Local Similarity 31.6%; Pred. No. 1.1e-35;
Matches 151; Conservative 84; Mismatches 178; Indels 65; Gaps 10;
QY 27 FLAFAAALGPLSGFALGVSSPAIPSLQRAAPPAPRLDDAAAFAGAVVTLGAAAGVL 86
Db 16 FFVCFLAALAGLLFGLDGVLAGLPFITDEF---QITATQEWVSSMFGAAVAVG 71
QY 87 GGWLVDVDRAGKLSLLCSVPVFAVITAAQDVWMLLGGRLTGLACGVASVAVPVLIS 146
Db 72 SGWLVSFLGRKKSMLMIGAILFVAGSLFSAAPNVEVLISRVLLGAVGVASYTAPLYLS 131
QY 147 BIAYPAVRGLGSCVQLMVMVVGILLAYLAGWVLE---WRWLAVLGVCPVPSLMLLMCFM 202
Db 132 BIAEKIRGSMISVQLMITIGLCAYLSDTAFYSVSGAWRWMLGVIIIPAILLIGVFFL 191
QY 203 PETRFLITQHRROEAMAL-----RFLWGEQGWEDDPIGAEOSFH 244
Db 192 PDSRPFVFAKRRFHAERVLLRDLTSAAKRELDEITRESLVQKQSGW----- 239
QY 245 LALLRQPGIYKPF1-IGVSLMAFQOLSGVNAVMAFETIPEAKFKDSS---LASVVGV 300

Db 240 -ALFKENSFRRAVFGILLQVWQOFTGMNVINYAPKIFELAGYINTTQMGWTVIVGL 298
QY 301 IQVLTFAVALIMDRAGRELLVLSGVWVFTSFAFGAYFKLTQGGPGNSHVAISAPVS 360
Db 299 TNVLATFIAIGLVDWRGRKPTLTGLFLVMA-----IGMGLTGMHIGIHS-P-S 346
QY 361 AQPVDASVGLAWLAVGSMCLFIAGFVAGWGPPIWLLMSEIFPLHVKGVATGICVLTNWLM 420
Db 347 AQ-----YFAIAMLMLFIIGFAMWAGPLWLCSEIQPLKGRDFGTCSTATTNWIA 397
QY 421 AFLVTKFSSLMVLRPYGAFWLASAFICFSPVLFTLCVPETKGTLEQITAH-FEGR 477
Db 398 NMIVGATFLTNLNNLGNANTFWYAGLVNLFILLTLWLPETKHSLEHLERNLMKGR 455

RESULT 15
G84864
Probable membrane transporter (imported) - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: G84864
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Varakan, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: AB4420; MUID:20083487; PMID:10617197
A;Accession: G84864
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-521 <STO>
A;Cross-references: GB:AE002093; NID:g2289003; PIDN:AAB64332.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g43330
A;Map position: 2
C;Superfamily: glucose transport protein

Query Match 23.7%; Score 582; DB 2; Length 521;
Best Local Similarity 31.0%; Pred. No. 1.6e-35;
Matches 147; Conservative 76; Mismatches 191; Indels 60; Gaps 9;
QY 33 AALGPLSGFALGVSSPAIPSL-----QRAAPPAPRLDDAAAFAGAVVTL---GAA 81
Db 38 AGIGLLDFGYDTGVISGALLYIKDDFEVVKQSFLQVYVNSSTSSKLETIVSMALVGM 97
QY 82 AGVGLGWLVDRAGKLSLLCSVPVFAVITAAQDVWMLLGGRLTGLACGVASVLA 141
Db 98 IGAAGGWINDYGRKKTATFADVWFAAGAVMAAAPDPVYLISGRLLVGLGVASVTA 157
QY 142 PVYISEIAYPAVRGLGSCVQLMVMVVGILLAYLAGWVLE---EWRWLAVLGVCPVPSLML 196
Db 158 PVYIAEASPSVREGGLVSTNVLMTTGGQFLSYLVNSAFTQVPGTWRMMLGVSGVPAVIQF 217
QY 197 LLMCFMPEIPREFLLTQHRROEAMALRFLWG-----SEQWEDDPIGAEQ 241
Db 218 LLMFNPESPRFLPMKRNKAEATQVLTARTYDLSREDEIDHLSAAEERKQKRTVG--- 274
QY 242 SFHLALLRQPGIYKPFITIGVSLMAFQOLSGVNAVMAFETIPEAKFKDSSLA---SVVV 298
Db 275 --YLDVFRSKELRLAFLAGAGLQAFQFTGINTVMYSPITVQWAGFHSNQLALFLSLIV 332
QY 299 GVTQVLETAVALIMDRAGRELLVLSGVWVFTSFAFGAYFKLTQGGPGNSHVAISAP 358
Db 333 AAMNAAGTVVGIYFIDHCHGRKKLAL-----SSLFVGIISL-----LILSVS 373
QY 359 VSAQPDASVG--LAWLAVGSMCLFIAGFVAGWGPPIWLLMSEIFPLHVKGVATGICVLT 416
Db 374 FFQKSETSSDGLYGVWLVGLALYIYFAPAGMGVFPVWVNSIYPOYRIGCGMGATV 433
QY 417 NWLMAFLVTKFSSLMVLRPYGAFWLASAFICFSPVLFTLCVPETKGTLEQI 470
Db 434 NWISNLIVAOFTLTIAEAGTGMTELLIAGLAVLAVIFVIVFVPEQTGLIFSEV 487

Search completed: September 27, 2004, 19:30:11
Job time : 51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 19:09:21 ; Search time 25 Seconds
(without alignments)
993.499 Million cell updates/sec

Title: US-09-886-954A-1

Perfect score: 2457

Sequence: 1 MTPEDPEETQPLLGPPGSA.....CVPETKGTLQITAHFGR 477

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2445	99.5	477	1	GTR8_HUMAN
2	2162	88.0	478	1	GTR8_RAT
3	2156	87.7	477	1	GTR8_MOUSE
4	1537.5	62.6	334	1	GTR8_BOVIN
5	948	38.6	507	1	GTR6_HUMAN
6	607	24.7	461	1	CSBC_BACSU
7	606.5	24.7	472	1	ARAE_KLEOX
8	603	24.5	472	1	ARAE_ECOLI
9	591.5	24.1	464	1	GLP_ECOLI
10	584.5	23.8	575	1	ITR1_SCHPO
11	579.5	23.6	494	1	GTR3_BOVIN
12	575.5	23.4	494	1	GTR3_SHEEP
13	574.5	23.4	493	1	GTR3_RAT
14	570.5	23.2	496	1	GTR3_HUMAN
15	569.5	23.2	493	1	GTR3_MOUSE
16	565.5	23.0	495	1	GTR3_CANFA
17	554.5	22.6	464	1	ARAE_BACSU
18	550	22.4	629	1	MYCT_HUMAN
19	546	22.2	457	1	XYLT_LACBR
20	546	22.2	491	1	XYLE_ECOLI
21	543	22.1	523	1	GTR2_MOUSE
22	542.5	22.1	522	1	GTR2_RAT
23	542	22.1	534	1	HUP3_CHLKE
24	536	21.8	509	1	GTR4_HUMAN
25	535.5	21.8	522	1	STP1_ARATH
26	531	21.6	523	1	STC_RICCO
27	530.5	21.6	468	1	GLCP_SYNY3
28	530	21.6	509	1	GTR4_RAT
29	529.5	21.6	540	1	HUP2_CHLKE
30	529	21.5	509	1	GTR4_MOUSE
31	528.5	21.5	534	1	HUP1_CHLKE
32	526.5	21.4	496	1	GTR3_CHICK
33	523	21.3	612	1	ITR2_YEAST

RESULT 1

34	520.5	21.2	492	1	GTR1_MOUSE
35	516	21.0	584	1	ITR1_YEAST
36	515.5	21.0	492	1	GTR1_BOVIN
37	514.5	20.9	492	1	GTR1_HUMAN
38	513.5	20.9	492	1	GTR1_RAT
39	513	20.9	524	1	GTR2_HUMAN
40	512.5	20.9	490	1	GTR1_CHICK
41	510.5	20.8	557	1	ITR2_SCHPO
42	507.5	20.7	492	1	GTR1_RABIT
43	507	20.6	400	1	GTR3_RABIT
44	499.5	20.3	473	1	GLF_ZYMO
45	498	20.3	533	1	GTR2_CHICK

ALIGNMENTS

GTR8_HUMAN STANDARD; PRT; 477 AA.

AC Q9NY64; Q9NSC4; (Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Solute carrier family 2, facilitated glucose transporter, member 8

DE (Glucose transporter type 8) (Glucose transporter type X1).

GN SLC2A8 OR GLUT8 OR GLUTX1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

PN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

FX MEDLINE=20283667; PubMed=10821868;

RA Doerge H., Schuermann A., Bahrenberg G., Brauers A., Joost H.-G.; "GLUT8, a novel member of the sugar transport facilitator family with glucose transport activity.";

RT J. Biol. Chem. 275:16275-16280(2000).

RL [2]

RN SEQUENCE FROM N.A.

RP MEDLINE=20138191; PubMed=10671487;

RX Ibberson M.R., Uldry M.A., Thorens B.; "GLUTX1, a novel mammalian glucose transporter expressed in the central nervous system and insulin-sensitive tissues.";

RT J. Biol. Chem. 275:4607-4612(2000).

RL [-] FUNCTION: Insulin-regulated facilitative glucose transporter. Binds cytochalasin B in a glucose-inhibitable manner. Seems to be a dual-specific sugar transporter as it is inhibitable by fructose (By similarity).

CC [-] SUBCELLULAR LOCATION: Integral membrane protein. Principally intracellular. May move between intracellular vesicles and the plasma membrane. The dileucine internalization motif is critical for intracellular sequestration (By similarity).

CC [-] TISSUE SPECIFICITY: Highly expressed in testis, but not in testicular carcinoma. Lower amounts present in most other tissues.

CC [-] INDUCTION: In testis, downregulated by estrogen.

CC [-] SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE TRANSPORTERS SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

CC EMBL; Y17801; CAB89809.1; --

CC EMBL; AJ245937; CAB75702.1; --

CC EMBL; HGNC:13812; SLC2A8.

CC MIM; 605245; --

CC GO; GO:0005887; C:integral to plasma membrane; TAS.

```

DR GO; GO:0005355; F:glucose transporter activity; TAS.
DR GO; GO:0005975; P:carbohydrate metabolism; TAS.
DR GO; GO:0015758; P:glucose transport; TAS.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF0083; sugar tr; 1.
DR PRINTS; PRO0171; SUGSTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Sugar transport; Transmembrane; Glycoprotein;
Multigene family.
FT DOMAIN 1 25 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 26 46
FT DOMAIN 47 70 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 71 91
FT DOMAIN 92 96 2 (POTENTIAL).
FT TRANSMEM 97 117 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 118 127 3 (POTENTIAL).
FT TRANSMEM 128 148 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 149 156 4 (POTENTIAL).
FT TRANSMEM 157 177 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 178 182 5 (POTENTIAL).
FT TRANSMEM 183 203 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 204 256 6 (POTENTIAL).
FT TRANSMEM 257 277 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 278 292 7 (POTENTIAL).
FT TRANSMEM 293 313 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 314 319 8 (POTENTIAL).
FT TRANSMEM 320 340 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 341 367 9 (POTENTIAL).
FT TRANSMEM 368 388 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 389 404 10 (POTENTIAL).
FT TRANSMEM 405 425 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 426 438 11 (POTENTIAL).
FT TRANSMEM 439 459 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 460 477 12 (POTENTIAL).
FT SITE 12 13 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 349 349 DILEUCINE INTERNALIZATION MOTIF (BY
FT CONFLICT 377 377 SIMILARITY).
FT CONFLICT 456 457 N-LINKED (GLCNAC. . .) (By similarity).
FT CONFLICT 462 462 S -> N (IN REF. 2).
FT SEQUENCE 477 AA; 50792 MW; 0B480F94B0AEE76 CRC64;
Query Match 99.5%; Score 2445; DB 1; Length 477;
Best Local Similarity 99.6%; Pred. No. 1.6e-164;
Matches 475; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MTPEDPETQLPGPGSGAPRRGVFLAFAAALGPLSGFALGYSPAPISLQRAAPP 60
Db 1 MTPEDPETQLPGPGSGAPRRGVFLAFAAALGPLSGFALGYSPAPISLQRAAPP 60
Qy 61 APRLDDAAASWFGAVVTLGAAAGVGLGWLVDRAGRKLSLLCSVPPFVAGFAVITAQDV 120
Db 61 APRLDDAAASWFGAVVTLGAAAGVGLGWLVDRAGRKLSLLCSVPPFVAGFAVITAQDV 120
Qy 121 WMLLGGRLTLGACVASLAPVYIIEIAPVAVRGLLGSVCQLMVMVVGILLAYLAGWLE 180
Db 121 WMLLGGRLTLGACVASLAPVYIIEIAPVAVRGLLGSVCQLMVMVVGILLAYLAGWLE 180
Qy 181 WRWLAVLGVCPSPSLMLLMCFMPTPRFLLTHRRQEAAMALFLMGSEOGWEDPPIGAE 240
Db 181 WRWLAVLGVCPSPSLMLLMCFMPTPRFLLTHRRQEAAMALFLMGSEOGWEDPPIGAE 240
Qy 241 QSFHLALLRPGYKPIIGVSLMAFQOLSGVNAVFYAEIIFEEAKFKDSSLASVVVGV 300
Db 241 QSFHLALLRPGYKPIIGVSLMAFQOLSGVNAVFYAEIIFEEAKFKDSSLASVVVGV 300
Qy 301 IQVLTFAVALIMDRAGRRLLLVLSGVVWFSTAFGAYFKLTQGGPGNSHVAISAPVS 360

```

RESULT 2

GTR8 RAT

ID GTR8 RAT STANDARD; PRT; 478 AA.

AC Q9JUZ1; Q9JUA6;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Solute carrier family 2, facilitated glucose transporter, member 8

DE (Glucose transporter type 8) (Glucose transporter type XI).

GN SLC2A8 OR GLUT8 OR GLUTX1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_taxID=10116;

RN [1]

SEQUENCE FROM N.A., CHARACTERIZATION, AND MUTAGENESIS OF

RP 12-LEU-LEU-13.

RC TISSUE=Testis;

RX MEDLINE=20138191; PubMed=10671487;

RA Ibberson M.R., Uldry M.A., Thorens B.;

RT "GLUTX1, a novel mammalian glucose transporter expressed in the

RL central nervous system and insulin-sensitive tissues.";

RN J. Biol. Chem. 275:4607-4612(2000).

RN [2]

SEQUENCE FROM N.A.

RC TISSUE=Testis;

RA Ishibashi K.;

RT "Molecular cloning of a new putative glucose transporter.";

RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

TISSUE SPECIFICITY.

MEDLINE=20283667; PubMed=10821868;

RX Dege H., Schuermann A., Bahrenberg C., Brauers A., Joost H.-G.;

RA "GLUT8, a novel member of the sugar transport facilitator family with

RT glucose transport activity.";

RL J. Biol. Chem. 275:16275-16280(2000).

CC -!- FUNCTION: Insulin-regulated facilitative glucose transporter.

CC Binds cytochalasin B in a glucose-inhibitable manner. Seems to be

CC a dual-specific sugar transporter as it is inhibitable by

CC fructose.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Principally

CC intracellular. May move between intracellular vesicles and the

CC plasma membrane. The dileucine internalization motif is critical

CC for intracellular sequestration.

CC -!- TISSUE SPECIFICITY: Highly expressed in adult and pubertal testis,

CC but not prepubertal testis. Moderate expression in hypothalamus,

CC cerebellum, brainstem, hippocampus, and adrenal gland. Lower

CC amounts present in most other tissues.

CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE

CC TRANSPORTERS SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AJ245935; CAB75729.1; -.

CC EMBL; AB033418; BAA94383.1; -.

CC DR

DR


```
DR InterPro: IPR005828; Sub transporter.
DR InterPro: IPR005829; Sug transporter.
DR Pfam: PF00083; sugar_tr; 1.
DR PROSITE: PS00850; MFS; 1.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; PARTIAL.
KW Transport; Sugar transport; Transmembrane; Multigene family.
FT NON_TER 1 1
FT DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 15 35 5 (POTENTIAL).
FT DOMAIN 36 38 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 39 59 6 (POTENTIAL).
FT DOMAIN 60 113 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 114 134 7 (POTENTIAL).
FT DOMAIN 135 149 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 150 170 8 (POTENTIAL).
FT DOMAIN 171 176 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 177 197 9 (POTENTIAL).
FT DOMAIN 198 224 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 225 245 10 (POTENTIAL).
FT DOMAIN 246 261 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 262 282 11 (POTENTIAL).
FT DOMAIN 283 295 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 296 316 12 (POTENTIAL).
FT DOMAIN 317 334 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 334 AA; 36699 MW; 0EE9B670ADAB71DD CRC64;

Query Match. 62.6%; Score 1537.5; DB 1; Length 334;
Best Local Similarity 87.7%; Pred. No. 5.4e-101;
Matches 293; Conservative 17; Mismatches 23; Indels 1; Gaps 1;

QY 145 ISRIAYPAVAGLGSCVOLMVVGVGILLAYLAGVLEWRLAVLGCVPSPSLMLLLMCMPPE 204
Db 1 ISRIAYPEVGLGSCVOLMVVGVGILLAYLAGVLEWRLAVLGCVPSPSLMLLLMCMPPE 60
QY 205 TPREFLTQHRQRQAMALRFLWGSEQGWEDPPIGAE-QSFHALLRDPGIVKPPPIGVSL 263
Db 61 TPREFLSQHQHEAMAMQFLGVAQWEEPPFGAQQHDFVAQLRPPGVKPPPIGSL 120
QY 264 MAFQQLSGVNAVVFYATIFEEAKFKDSSLASVVGVQVLFATAVALINDRAGRRLLV 323
Db 121 MAFQQLSGVNAVVFYATIFEEAKFKDSSLASVVGVQVLFATAVALINDRAGRRLLT 180
QY 324 LSGVWVWVSTSAFAYEKLFGGPGNSHVAISAPVQAQPVDSVGLAWLAVGSMCLFIA 383
Db 181 LSGVWVWVSTSAFAYEKLFGGPGNSHVDLPALVSMEEAADTNVGLAWLAVGSMCLFIA 240
QY 384 GFVAGWGPPIPLLMSLFIPLHVKGVATGICVLTNWLMFLVTKFSSLMELVLRPYGAFWL 443
Db 241 GFVAGWGPPIPLLMSLFIPLHVKGVATGVCVLTNWLMFLVTKFSSLMELVLRPYGAFWL 300
QY 444 ASAFICFISVLFTLPCVPETKGTLEQITAHFEGR 477
Db 301 ASAFICFISVLFTLPCVPETKGTLEQITAHFEGR 334

RESULT 5
GTR6_HUMAN STANDARD; PRT; 507 AA.
AC Q8UGQ3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 6
DE (Glucose transporter type 6) (Glucose transporter type 9).
GN SLC2A6 OR GLUT9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukocyte;
```


DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR005829; Sug_transporter.
 DR InterPro; IPR003663; Sugar_transpt.
 DR Pfam; PF00083; sugar_tr; 1.
 DR PRINTS; PR00171; SUGTRNSPORT.
 DR TIGRFAMs; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transprot; Sugar transport; Transmembrane; Inner membrane; Symport;
 KW Complete proteome.
 FT DOMAIN 1 29 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 30 50 1 (POTENTIAL).
 FT DOMAIN 51 63 PERIPLASMIC (POTENTIAL).
 FT TRANSSEM 64 84 2 (POTENTIAL).
 FT DOMAIN 85 91 3 (POTENTIAL).
 FT TRANSSEM 92 112 4 (POTENTIAL).
 FT DOMAIN 113 114 PERIPLASMIC (POTENTIAL).
 FT TRANSSEM 115 135 5 (POTENTIAL).
 FT DOMAIN 136 154 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 155 175 6 (POTENTIAL).
 FT DOMAIN 176 178 PERIPLASMIC (POTENTIAL).
 FT TRANSSEM 179 199 7 (POTENTIAL).
 FT DOMAIN 200 257 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 258 278 8 (POTENTIAL).
 FT DOMAIN 279 297 PERIPLASMIC (POTENTIAL).
 FT TRANSSEM 298 318 9 (POTENTIAL).
 FT DOMAIN 319 325 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 326 346 10 (POTENTIAL).
 FT DOMAIN 347 361 PERIPLASMIC (POTENTIAL).
 FT TRANSSEM 362 382 11 (POTENTIAL).
 FT DOMAIN 383 404 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 405 425 12 (POTENTIAL).
 FT DOMAIN 426 427 PERIPLASMIC (POTENTIAL).
 FT TRANSSEM 428 446 12 (POTENTIAL).
 FT DOMAIN 447 472 CYTOPLASMIC (POTENTIAL).
 FT CONFLICT 26 28 SVA -> YDR (IN REF. 6).
 SQ SEQUENCE 472 AA; 51684 MW; 411990A441D44393 CRC64;

Query Match 24.5%; Score 603; DB 1; Length 472;
 Best Local Similarity 33.3%; Pred. No. 2.9e-35;
 Matches 158; Conservative 84; Mismatches 197; Indels 36; Gaps 9;

QY 9 TQPLLPGGSAAPGRVFLAAFAALGPLSGFALGVSSPAISLQRAAPPAPRLDAA 68
 DB 6 TESALTFRSLDTRMMNFVSV-AAVAGLGLFDGIVAGALPFTDHFVLTSLRQEE-- 62
 QY 69 ASWFGAVVTLGAAGGVGLGVLDRAGRKLSLLCSVPFVAGFAVITAADQVMMLLGRL 128
 DB 63 --WVSSMMLGAALGALFNGWLSFRLGRKYSLSMAGAILFVLSGSFAFATSVEMLIAARV 120
 QY 129 LTGLACGVASIVAPVYISETIAPVAVRGLLSCVQLMVVVGILLAYLAGWVLEW--RWLAV 186
 DB 121 VLGIAGVIASTAPLYLSEKASENVRGKMSYQMLVTLGLVLAFLSDTAFSYSGNWRAM 180
 QY 187 LG--CVPEPSMLLLMCFMPTETPRLLTCHRRQEAAMALRLWLGSEQGWEDPPIGAEOSFH 244
 DB 181 LGVIALPAVLLIILVFLDPSRWLAEGRIHIEAEVLRMLDTSKAREELNEIRSLK 240
 QY 245 L-----ALLR-QPGIYKPFIIIGVSLMAFQQLSGVNAVVFYAEITFEFAKF---KDSIAS 295
 DB 241 LKQSGMALFKINVRRAVFLGMLLQAMQQTGNIIYVYAPRIPKVGAGFTTTEQQMIAT 300
 QY 296 VVUGVIVQLFTAVAAALMDRAGRLLLVLSGVVWVFTSAPGAVFKUTQGGFGNSHVAI 355
 DB 301 LVVGLTFEMFAFIAVFTVDKAGRPALKIGFSVMALGTLVLG--YCLMQDFNGT----- 352
 QY 356 SAPVSAQPDVASGLAWLAVGSMCLFIAGFVAGVGPIPLWLMSEIFPLHWKGVATGICVL 415
 DB 353 -----ASSGLSMLSVGTMWMCITAGYAMSAAPVVMILCSIQPLKCRDFGICSTT 402
 QY 416 TNWLMFAVLTKEFSSLMELVRPYGAFWLASAFICFVSLFTLFCVPETKGGKTLEQI 470

Db 403 TNWVSNMIGATFLTLDSIGAGCTFWLTALNIAFVGITFWLIPETKNVTLHI 457

RESULT 9
 GALP_ECOLI STANDARD; PRT; 464 AA.
 AC P37021;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Galactose-proton symporter (Galactose transporter).
 GN GALP OR B2943 OR C3529.
 OS Escherichia coli.
 OS Escherichia coli. O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562, 217992;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RL Roberts P.E.;
 RL Thesis (1992), University of Cambridge, U.K.
 RN [2]
 RP SEQUENCE FROM N.A.
 RL STRAIN=K12 / MG1655;
 RL MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R.; Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RA "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474 (1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RL STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RL MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackert J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
 CC -!- FUNCTION: UPTAKE OF GALACTOSE ACROSS THE BOUNDARY MEMBRANE WITH
 THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -!- SIMILARITY: Belongs to the sugar transporter family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U28377; AAA69110.1; .
 DR EMBL; AE000377; AAC75980.1; .
 DR EMBL; AE016766; AA881977.1; ALT_INIT.
 DR PIR; F65079; F65079.
 DR EcoGene; EG12148; galp.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR005829; Sug_transporter.
 DR InterPro; IPR03663; Sugar_transpt.
 DR Pfam; PF00083; sugar_tr; 1.
 DR PRINTS; PR00171; SUGTRNSPORT.
 DR TIGRFAMs; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transprot; Sugar transport; Transmembrane; Inner membrane; Symport;
 KW Complete proteome.
 FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).

```
FT TRANSMEM 16 36 1 (POTENTIAL).
FT DOMAIN 37 56 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 57 77 2 (POTENTIAL).
FT DOMAIN 78 84 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 85 102 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 106 112 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 113 133 4 (POTENTIAL).
FT DOMAIN 134 139 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 140 160 5 (POTENTIAL).
FT DOMAIN 161 171 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 172 192 6 (POTENTIAL).
FT DOMAIN 193 250 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 251 271 7 (POTENTIAL).
FT DOMAIN 272 290 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 291 311 8 (POTENTIAL).
FT DOMAIN 312 321 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 322 342 9 (POTENTIAL).
FT DOMAIN 343 351 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 352 372 10 (POTENTIAL).
FT DOMAIN 373 394 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 395 415 11 (POTENTIAL).
FT DOMAIN 416 437 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 438 464 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 464 AA; 50982 MW; 07E08935BD8E3F8E CRC64;

Query Match 24.1%; Score 591.5; DB 1; Length 464;
Best Local Similarity 32.4%; Pred. No. 1.8e-34;
Matches 155; Conservative 83; Mismatches 175; Indels 65; Gaps 11;

QY 27 FLAFAAALGPLSGFALGYSIPAISQRAAPPAPRDLDDAAAFGAVVLTGAAGVGL 86
Db 16 PFVCFALAAGLLGDLGIVTALGALPFTADEP---QITSTQEWVSMFEGAAGVG 71

QY 87 GGWLVDRAGRKLSLLCSVPFVAGFAVITAQDVNMLLGRLLTGLACGVASLAPVVIS 146
Db 72 SGWLSFKLGRKSLMIGAILFVAGSLFSAAPNVEVLILSRVLLGLAVGASVYAPLVLS 131

QY 147 EIAYPAVRLGSCVOLMVVVGILLAYLA---GWVLEWRMLAVLGCVPSPSLMLLMCFM 202
Db 132 EIAEPKIRGSMISYQMLMITIGILGAYLSDTAFSYTGAWRWMLGVIIIPAILLLIGVFFL 191

QY 203 PETRFLTLQHRROEAMAL-----RFLWGSQQGWEDPPIGAEQSFH 244
Db 192 PDSPRFWFAKRFVDAERVLRLRDTSAEKHELDEIRSLQVKSGM----- 239

QY 245 LALLRQGIYKPII-IGVSLMAFQOLSGVNAVVFYAEITPEAKPKDSS---LASVVGV 300
Db 240 -ALFKENSNFRVFLGVLLQVMQOFTGMNVMYAPKIFELAGYTNTEQMWGVTVGL 298

QY 301 IQVLTFAAALIMDRAGRLLLVLSGVVMVFTSAPGAYFKLTQGGPGNSHVAISAPVS 360
Db 299 TNVLAFTAILGVDRWGRKPTLTIGFLVW-----AAG-----MGVLGTMHGHISPS-S 346

QY 361 AQPVDASVGLAWLVGSMCLFIAGPAVGWGPPIPLLMSIEIFPLHVKVATGICVLTNLM 420
Db 347 AQ-----YFAIALLMFIIVGVFAMSAGPLIWLVCSEIQPLKGRDFGICSTATNIA 397

QY 421 AFLVTKFSSLMELVRLPGAFWLASAPCIFSVLFTLCVPETPKGTILQITAH-PEG 477
Db 398 NMIVGATFLMTLNTLGNANTFWVYAAALNVLFILLTLVLPETPKHVSLEHIERNLKGR 455

RESULT 10
ITRI SCHPO STANDARD; PRT; 575 AA.
AC Q10286;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myo-inositol transporter 1.
GN ITRI OR SPAC4F8.15 OR SPAC7D4.01.
OS Schizosaccharomyces pombe (Fission yeast).
```

```
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=968 h90;
RX MEDLINE=98228265; PubMed=9560432;
RA Niederberger C., Graub R., Schweingruber A.-M., Fankhauser H.,
RA Rusu M., Poitelea M., Edeharter L., Schweingruber M.E.;
RT "Exogenous inositol and genes responsible for inositol transport are
RL required for mating and sporulation in Schizosaccharomyces pombe.";
Curr. Genet. 33:255-261(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gidley S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tavey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaue V., Motier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey P., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McComb W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
Nature 415:871-880(2002).
RC -!- FUNCTION: TRANSPORTER FOR MYO-INOSITOL.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the sugar transporter family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X98622; CAA67211.1; -.
DR EMBL; Z98530; CAB11061.1; -.
DR EMBL; Z99332; CAB16718.1; -.
DR F01; T43400; T43400.
DR GeneDB SPombe; SPAC4F8.15; -.
DR GO; GO:0000747; P:conjugation with cellular fusion; ISS.
DR GO; GO:0006629; P:lipid metabolism; ISS.
DR GO; GO:0007165; P:signal transduction; ISS.
DR GO; GO:0007151; P:sporulation (sensu Saccharomycetes); ISS.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transp.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGTNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
```



```
FT DOMAIN 385 399 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 400 420 11 (POTENTIAL).
FT DOMAIN 421 424 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 425 445 12 (POTENTIAL).
FT DOMAIN 446 494 CYTOPLASMIC (POTENTIAL).
FT SITE 277 279 DEFINES SUBSTRATE SPECIFICITY (BY SIMILARITY).
FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 358 358 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 385 389 PWFIV -> SLVYC (IN REF. 2).
SQ SEQUENCE 494 AA; EC42A3C648CAD23C CRC64;

Query Match 23.6%; Score 579.5; DB 1; Length 494;
Best Local Similarity 31.2%; Pred No. 1.4e-33;
Matches 157; Conservative 96; Mismatches 177; Indels 73; Gaps 14;

QY 13 LGPPGSGAPRRVFLAFAAALGSLGFGALGY--SSPAI-----PSLQRAAPPAPRL 64
Db 1 MGTTKVTAP---LIFAI SVATIGSFQGYNTGVINAPEAIIKDFLNYTLBERSEPPSS 56
QY 65 DDAASWFGAVVTLAGAGGVLG---GWLVDRAGRKSLLLCSVPFVAGFAVI---TAA 117
Db 57 VLLTSLWLSVAIF--SVGGMIGSFVGLFVNRFGNSMLIVNLLAAGCLMGFCXIA 114
QY 118 QDVMLLGLRLITGLACGVASIVAPVYISEIAYPAVRGLLSCVOLMVVVGILLAYLAGW 177
Db 115 ESVENLILGRLLIGLFCGLCTGFVPMYIGETSPALRGAFGLNQLGIVIGILVAQIFGL 174
QY 178 VL-----EWRWLAVLG--CVPSLMILLMCFMPTPRFLLTQHRQF-AMAAIRFLWSEQ 230
Db 175 KVIITGDLWPLLGLFTILPAIIQCAALPFCPSPRFLINRKEEKAKIQLRLWGTD 234
QY 231 GWEDPPIGAQSFH-----LALLQPGIYKPIITGVSLMAFQQLSGVNVNMFYAE 281
Db 235 VAQDIQEMKDSMRMSQKQVTVLELFRAPNVRQPIIISIMLQSLQSGINAVFYSTG 294
QY 282 IFEEAKFKDSSLASVVVGVIQVLTFAVAALIMDRAGRLILLVLSGVWVVFSTSAFGAYFK 341
Db 295 IFKDAVQEPVATIGAGVNTIFTVSVFLVERAGRTLHLI----- 337
QY 342 LTQGGPGNSHVAISAPVSAQPVDSVGLAVGSMCLFIAGFVAGVGPIPLWLMSEIF 401
Db 338 ----GLGGMAFCSILMTSLLLKDNYSWMSFCIGAILVFAFFBPGIPWFIABELF 393
QY 402 PLHVKGATGICVLNMLMARLVTKFSSLMELVLPYGAFL-ASAFICSY-----LF 454
Db 394 GGGPPAAMAVAGCSNWTSNFLV-----GLLPFAAFYLGAYVFIIVTVFLVIFWVF 445
QY 455 TLFCVPETKGTLEQITAHFEGR 477
Db 446 TFFKVPETGRTEFEITRAFEQ 468
```

RESULT 12

```
ID_GTR3 SHEEP STANDARD; PRT; 494 AA.
AC P47843.
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 3
GN (Glucose transporter type 3, brain).
DN SLC2A3 OR GLUT3 OR GLUT-3.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=Coopworth;
RX MEDLINE=96109471; PubMed=8653093;
RA Bennett B.L., Prosser C.G., Grigor M.R.;
```

```
RT "Isolation of cDNAs and tissue specific expression of ovine glucose
RT transporters.";
RL Biochem. Mol. Biol. Int. 37:9-16(1995).
RN [2]
RC SEQUENCE OF 88-248 FROM N.A.
RX TISSUE=Placenta;
RX MEDLINE=97392487; PubMed=9250701;
RA Currie M.J., Bassett N.S., Gluckman P.D.;
RT "Ovine glucose transporter-1 and -3: cDNA partial sequences and
RT developmental gene expression in the placenta.";
RL Placenta 18:393-401(1997).
CC -!- FUNCTION: Facilitative glucose transporter. Probably a neuronal
CC glucose transporter.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Highly expressed in brain.
CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose
CC transporter subfamily.
CC -----
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

EMBL; L39214; AAC41629.1; -;
EMBL; U89030; AAB49313.1; -;
InterPro; IPR007114; MFS.
InterPro; IPR005828; Sub transporter.
InterPro; IPR005829; Sug transporter.
InterPro; IPR003663; Sugar transprt.
Pfam; PF00083; sugactr; 1;
PRINTS; PR00171; SUGTRNSPORT.
TIGRFAMS; TIGR00879; SP; 1.
PROSITE; PS00850; MFS; 1.
PROSITE; PS00216; SUGAR TRANSPORT 1; 1.
PROSITE; PS00217; SUGAR TRANSPORT 2; 1.
Transmembrane; Sugar transport; Transport; Glycoprotein;
Multigene family.

```
FT DOMAIN 1 5 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 6 26 1 (POTENTIAL).
FT DOMAIN 27 64 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 65 85 2 (POTENTIAL).
FT DOMAIN 86 93 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 94 114 3 (POTENTIAL).
FT DOMAIN 115 124 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 125 145 4 (POTENTIAL).
FT DOMAIN 146 153 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 154 174 5 (POTENTIAL).
FT DOMAIN 175 183 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 184 204 6 (POTENTIAL).
FT DOMAIN 205 269 7 (POTENTIAL).
FT TRANSMEM 270 290 8 (POTENTIAL).
FT DOMAIN 291 304 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 305 325 9 (POTENTIAL).
FT DOMAIN 326 334 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 335 355 10 (POTENTIAL).
FT DOMAIN 356 363 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 364 384 11 (POTENTIAL).
FT DOMAIN 385 399 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 400 420 12 (POTENTIAL).
FT DOMAIN 421 424 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 425 445 13 (POTENTIAL).
FT DOMAIN 446 494 CYTOPLASMIC (POTENTIAL).
FT SITE 277 279 DEFINES SUBSTRATE SPECIFICITY (BY SIMILARITY).
FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 494 AA; A89204D3EA74BFBA CRC64;
```

Query Match 23.4%; Score 575.5; DB 1; Length 494;
Best Local Similarity 30.7%; Pred. No. 2.6e-33;
Matches 150; Conservative 94; Mismatches 179; Indels 65; Gaps 10;

QY 26 VFLAFAAALGPLSPFALGYSPAIPLSL-----QRAAPPAPRLDDAASWFGAV 75
 Db 10 LIFAISTATIGTSFGYNTGVINAPEAIIKDFLNYLLEERSETPPSSVLITLSWSLVAI 69
 QY 76 VTLGAAAGVGLGWLVDRAGRKLSLLCSVPFVAGPAVI--TAAQDVMMLLGRLITGL 132
 Db 70 FSVGGMIGSFVGLFVNRFGRRMLIVNLLAAGGLMGFCFKIAESVEMLIIGRLIIGL 129
 QY 133 ACGVASLAPVYIETAYPAVRGLGSCVQLVMVVGVGILLAYLAGWVL-----EWRWLAVLG 188
 Db 130 FCGLCTGFVPMYIGISPTALRGAFGLNQLGIVIGILVAQIFGLKVLCTEDLWPLLLG 189
 QY 189 --CVPSLMLLLCMFETPRFILLTOHRROR-AMAAIRFLWSEQWEDPPIGAEQSFH- 244
 Db 190 FTILPAIQCAALPFCESPRFLINKKEBEKAKELLORLWGTEDVAQDIQEMKDESMRM 249
 QY 245 -----LALLRPGYKPFPIIGVSLMAFQQLSGVNAVVFYAEIIFEEAKFKDSSLASV 296
 Db 250 SQEQVTVLELFRAPNYRQPIIISIMQLSQQLSGINAVFYSTGIPKDGAVQEPVYATI 309
 QY 297 VVGVIQVLFATAALIMDRAGRLLLVLSGVWVVFSTSAFGYFKLTQGGPGNSSHVAIS 356
 Db 310 GAGVWNTIFTWVSVELVERAGRTHLI-----GLGGWAFCSIL 348
 QY 357 APVSAQPVDSVGLAWLAVGSMCLFIAGFAVGMGPPIELMLSETFPLHVKGVAIGCVLT 416
 Db 349 MTISLLKKNYSWMSFCIGAILVFAFFGIPGPIEFWIVAEFLFGGPPAPMAVAGCS 408
 QY 417 NWLMFAFLVTKFSSLMVLPYGAFLW-ASAFICFSV-----LFTFLCVPETKGTLEQ 469
 Db 409 NWTSNELV-----GLLPSATFYLGAVFVFTVFLVIFWVTFEKFVETGRIFEE 460
 QY 470 ITAHFEGR 477
 Db 461 ITRAFEGQ 468

RESULT 13
 GTR3 RAT STANDARD; PRT; 493 AA.
 AC Q07647; Q62729;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Solute carrier family 2, facilitated glucose transporter, member 3
 DE (Glucose transporter type 3, brain).
 GN SLC2A3 OR GLUT3 OR GLUT-3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=94063043; PubMed=8243635;
 RA Nagamatsu S., Sawa H., Kanada K., Nakamichi Y., Yoshimoto K.,
 RA Hoshino T.;
 RT "Neuron-specific glucose transporter (NSGT): CNS distribution of
 RT GLUT3 rat glucose transporter (RGT3) in rat central neurons.";
 RL FEBS Lett. 334:289-295(1993).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=96042338; PubMed=7475896;
 RA Krishnan S.N., Haddad G.G.;
 RT "Cloning of glucose transporter-3 (GLUT3) cDNA from rat brain.";
 RL Life Sci. 56:1193-1197(1995).
 CC -!- FUNCTION: Facilitative glucose transporter. Probably a neuronal
 CC glucose transporter.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: Highly expressed in brain.
 CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose
 CC transporter subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; DI3962; BAA03065.1; -;
 DR EMBL; U17978; AAA62503.1; -;
 DR PIR; S38981; S38981.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR005829; Sug_transporter.
 DR InterPro; IPR003663; Sugar_transp.
 DR Pfam; PF00083; sugar tr; 1.
 DR PRINTS; PR00171; SUGTRNSPORT.
 DR TIGRFAMS; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 DR Transmembrane; Sugar transport; Transport; Glycoprotein;
 KW Multigene family.
 FT DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 11 31 1 (POTENTIAL).
 FT DOMAIN 32 64 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 65 85 2 (POTENTIAL).
 FT DOMAIN 86 93 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 94 114 3 (POTENTIAL).
 FT DOMAIN 115 124 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 125 145 4 (POTENTIAL).
 FT DOMAIN 146 153 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 154 174 5 (POTENTIAL).
 FT DOMAIN 175 183 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 184 204 6 (POTENTIAL).
 FT DOMAIN 205 269 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 270 290 7 (POTENTIAL).
 FT DOMAIN 291 304 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 305 325 8 (POTENTIAL).
 FT DOMAIN 326 334 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 335 355 9 (POTENTIAL).
 FT DOMAIN 356 363 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 364 384 10 (POTENTIAL).
 FT DOMAIN 385 399 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 400 420 11 (POTENTIAL).
 FT DOMAIN 421 424 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 425 445 12 (POTENTIAL).
 FT DOMAIN 446 493 CYTOPLASMIC (POTENTIAL).
 FT SITE 277 279 DEFINES SUBSTRATE SPECIFICITY (BY
 FT SIMILARITY).
 FT CARBOHYD 43 43 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 55 55 R -> S (IN REF. 2).
 FT CONFLICT 447 447 S -> F (IN REF. 2).
 FT CONFLICT 470 470 H -> Q (IN REF. 2).
 SQ SEQUENCE 493 AA; 53580 MW; 745446B59BDF4399 CRC64;
 Query Match 23.4%; Score 574.5; DB 1; Length 493;
 Best Local Similarity 31.1%; Pred. No. 3e-33;
 Matches 156; Conservative 93; Mismatches 176; Indels 77; Gaps 15;
 QY 18 GSAPRGRRVFLAAFAAALGSLPFGFALGYSSPAISLQRAAPPA-----PRLDD 66
 Db 2 GTAKVTPSLVFAVTVATIGSFQGYNTGVIN-----APETIIKDFLNYLLEERLED 52
 QY 67 AAASWFGAVTVLGA-----AAGVLG----GMLVDRAGRKLSLLGCVSPFV-----AGF 111
 Db 53 LPRE--GLLITLWSLCAVIFSVGMIGSFVGLFVNFRGRNMLNLLNLAAILGGCLMGF 110
 QY 112 AVITAAQDVMMLLGGRLLITGLACGVASLVAIPVYIETAYPAVRGLGSCVQLVMVVGILL 171
 Db 111 AKI--AESVEMLIIGRLIIGIFCGLCTGFVPMYIGVSPALRGAFGLNQLGIVGILV 168


```
FT CONFLICT 357 358 DN -> NH (IN REF. 3).
FT CONFLICT 376 376 F -> C (IN REF. 3).
FT CONFLICT 425 425 H -> Y (IN REF. 3).
FT CONFLICT 432 432 H -> L (IN REF. 3).
FT CONFLICT 482 482 E -> G (IN REF. 3).
SQ SEQUENCE 496 AA; 53924 MW; F601CD6892F16516 CRC64;

Query Match 23.2%; Score 570.5; DB 1; Length 496;
Best Local Similarity 31.6%; Pred. No. 5.8e-33;
Matches 152; Conservative 88; Mismatches 190; Indels 51; Gaps 9;

Qy 26 VFLAFAAALGPLSGFGALGSSPAISLQOR-----AAPPAPSLDDAAASWFCV 75
Db 10 LIPALTATIGSFQGYNTGVINAPKIKKIFINKLTLDKGNAPSEVLLTSLMSLVAI 69
Qy 76 VTLGAAGVGLGWLVDRAKRLSLLCSVPFVAG---FAVITRAQDVMWLLGRLITGL 132
Db 70 FSVGGMIGSFVGLFVNRFGRRNSMLIVNLLAVTGCGFMGLCKVAKSVEMILGLRVLGL 129
Qy 133 ACGVASLVAPVYISIAIYPAVRGLLGSQVQLMVVVGILLVLAG--WVL--EWELAVLG 188
Db 130 FCGLCTGFVNYIGIISPTLRGAFGLNQGLGVVGLVAGIFGLFELGSEELWPLLGL 189
Qy 189 --CVPELSMLLLCMFETPRFLTLQHRROE-AVAALRFLMGSEQGWEDPPIGAEQSPH- 244
Db 190 FTILPAILQSAALPFCSPESPRFLINRKEENAKQLQLRWGTQDVSDIQEMKDSARM 249
Qy 245 -----LALLRPGYIKPIIGVSLMAFQOLSGVNAVMPYAEITPEAKFDSSIASV 296
Db 250 SQEKQVTLVLEFRVSSVRQPIIISIVLQSLQSCINAVFYSTGIFKADGAVQEPVATI 309
Qy 297 VVGVIQVLTAVAAALMDRAGRRLLLVLSGVVVFSTAFGAYFKLTGGGNSHVAIS 356
Db 310 GAGVNTITFVVSFLVERAGRTLHMI-----GUGGNAPFCSTL 348
Qy 357 APVSAQPVDSVGLAWLAVGSMCLFIAGFVAGVGWGPWPMLLMSEIFFPLHVKVATGICVLT 416
Db 349 MTVSLLLKDNVNGMSFVCIGAILVFAFFEGPGIPWFIIVAEILFSGQPRPAANAVAGCS 408
Qy 417 NWLMFVLTKFESSLMEVLRYPGFVWLASAFICISVLTFLFCVETPKTKTLEQITAHFEG 476
Db 409 NWTNLFVGLLFPSSAAHYLGAY-VFIPTGFLITFLAFTPKVPETRGTPEDITRAFEG 467
Qy 477 R 477
Db 468 Q 468

RESULT 15
GTR3 MOUSE
ID_GTR3_MOUSE STANDARD; PRT; 493 AA.
AC P32037;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 3
DE (Glucose transporter type 3, brain).
DE SLC2A3 OR GLUT3 OR GLUT-3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC MEDLINE=92112695; PubMed=1730609;
RX Nagamatsu S., Kornhauser J.M., Seino S., Mayo K.E., Steiner D.F.,
RA Bell G.I.
RT "Glucose transporter expression in brain. cDNA sequence of mouse
RT GLUT3, the brain facilitative glucose transporter isoform, and
RT identification of sites of expression by in situ hybridization."
RL J. Biol. Chem. 267:467-472(1992).
RN [2]
```

```
RP SEQUENCE FROM N.A.
RA Takeda J., Minokoshi Y., Yasuda K., Kayano T., Graeme B.I.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Facilitative glucose transporter. Probably a neuronal
CC glucose transporter.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Highly expressed in brain.
CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose
CC transporter subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M75135; AAA37704.1; -
DR EMBL; X61093; CAA43406.1; -
DR EMBL; U11853; AAB60666.1; -
DR EMBL; U11844; AAB60666.1; JOINED.
DR EMBL; U11845; AAB60666.1; JOINED.
DR EMBL; U11846; AAB60666.1; JOINED.
DR EMBL; U11848; AAB60666.1; JOINED.
DR EMBL; U11849; AAB60666.1; JOINED.
DR EMBL; U11850; AAB60666.1; JOINED.
DR EMBL; U11851; AAB60666.1; JOINED.
DR EMBL; BC034122; AAH34122.1; -
DR EMBL; BC058811; AAH58811.1; -
DR PIR; A41751; A41751.
DR MGD; MGI:95757; SLC2a3.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR005829; Sug transporter.
DR InterPro; IPR003663; Sugar_transp.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transporter; Transport; Glycoprotein;
KW Multigene family.
FT DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 11 31 1 (POTENTIAL).
FT DOMAIN 32 64 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 65 85 2 (POTENTIAL).
```

Search completed: September 27, 2004, 19:27:12
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 28, 2004, 00:53:58 : Search time 4589 Seconds
(without alignments)
4505.258 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPEDPETQPLLPGGSA.....CVPETKGTLEQITAFEGR 477

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0986954/runat_27092004.170631_13388/app_query.fasta_1.647
-DB=GenEmbl -QMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=ptlo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0986954 -CGN 1_5600 @runat_27092004.170631_13388 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

- 1: gb.pa.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.or.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*

- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.man.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	2457	100.0	2080	6	AX179740	Sequence
2	2457	100.0	2145	9	BC019043	BC019043 Homo sapi
3	2448	99.6	1873	9	HSA245937	AJ245937 Homo sapi
4	2448	99.6	2217	6	AX076667	AX076667 Sequence
5	2445	99.5	1508	9	HSA17801	Y17801 Homo sapien
6	2162.5	88.0	2087	6	AX076669	AX076669 Sequence
7	2162.5	88.0	2087	10	RNO245935	AJ245935 Rattus no
8	2159	87.9	1490	10	MMU17802	Y17802 Mus musculu
9	2156	87.7	1843	10	AF232061	AF232061 Mus muscu
10	2150	87.5	2072	6	AX076671	AX076671 Sequence
11	2150	87.5	2072	10	MMU245936	AJ245936 Mus muscu
12	2146	87.3	2189	10	AB033418	AB033418 Rattus no
13	1554.5	63.3	1461	6	AX191507	AX191507 Sequence
14	1537.5	62.6	1012	4	AF321324	AF321324 Bos tauru
15	1411.5	57.4	1929	5	BC049409	BC049409 Danio rer
16	1407	57.3	1449	5	AB083371	AB083371 Gallus ga
17	1268	51.6	789	6	AX191497	AX191497 Sequence
18	1167	47.5	1493	10	BC021758	BC021758 Mus muscu
19	1135.5	45.8	738	4	AF495799	AF495799 Ovis arie
20	972.5	39.6	2011	6	AX076679	AX076679 Sequence
21	958.5	39.0	1580	9	HSA17803	Y17803 Homo sapien
22	958.5	39.0	2356	6	AX677165	AX677165 Sequence
23	958.5	39.0	2356	9	BC013740	BC013740 Homo sapi
24	958.5	39.0	2390	6	BD127367	BD127367 Primer fo
25	958.5	39.0	2390	9	AK074927	AK074927 Homo sapi
26	958.5	39.0	2487	9	HSA011372	AJ011372 Homo sapi
27	949	38.6	1541	6	AX076677	AX076677 Sequence
28	910.5	37.1	2364	10	BC058210	BC058210 Mus muscu
29	904.5	36.8	2171	6	BD127313	BD127313 Primer fo
30	904.5	36.8	2171	9	AK074836	AK074836 Homo sapi
31	891.5	36.3	9582	10	MMU413951	AJ413951 Mus muscu
32	891.5	36.3	198006	9	AL731852	AL731852 Mus muscu
33	870.5	35.4	225370	9	AL445222	AL445222 Human DNA
34	867.5	35.3	3114	9	AK096630	AK096630 Homo sapi
35	824.5	33.6	250940	2	AC096489	AC096489 Rattus no
36	807	32.8	241623	2	AC103008	AC103008 Rattus no
37	804.5	32.7	1673	3	AY122079	AY122079 Drosophil
38	789.5	32.1	3160	3	BT003466	BT003466 Drosophil
39	788.5	32.1	2093	3	BT010286	BT010286 Drosophil
40	788	32.1	2937	3	AY119564	AY119564 Drosophil
41	736	30.0	1557	8	AY093274	AY093274 Arabidops
42	736	30.0	1866	8	AY059848	AY059848 Arabidops
43	734	29.9	1464	8	AY124845	AY124845 Arabidops
44	734	29.9	1587	3	AY058435	AY058435 Drosophil
45	734	29.9	1695	8	AF412060	AF412060 Arabidops

ALIGNMENTS

RESULT 1

```

AX179740
LOCUS AX179740 2080 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 28 from Patent WO0146258.
ACCESSION AX179740
VERSION AX179740.1 GI:15132104
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS Baughn,M.R., Burford,N., Au-Young,J., Lu,D.A., Yang,J., Reddy,R.,
Lal,P., Hallman,J.L., Acimzai,Y., Yue,H., Nguyen,D.B., Yao,M.G.,
Gandhi,A.B., Tang,Y.T. and Khan,F.A.
TITLE Transporters and ion channels
JOURNAL Patent: WO 0146258-A 28 JUN-2001;
FEATURES
source Location/Qualifiers
1..2080
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 1416107CB1"
ORIGIN
Alignment Scores:
Pred. No.: 2 448-167 Length: 2080
Score: 2457.00 Matches: 477
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-886-954A-1 (1-477) x AX179740 (1-2080)
QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db 46 ATGACGCCGAGAGCCAGAGAAACCCAGCGCTTCCTGGGGCCCTCTGGGGCGAGCGG 105
QY 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
Db 106 CCCCAGCGCGCGCGCTTCCTCGCGCGCTTCGCGCGCTGCGCGCGCTGCGCGCGCTC 165
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaPro 60
Db 166 GGTTCGCGCTCGGTACAGCTCCCGGCCATCCCTAGCTGACGCGCGCGCGCGCGCG 225
QY 61 AlaProArgLeuAspAlaAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGly 80
Db 226 GCGCGCGCGCTGGAGCGCGCGCTCCTCGTTCGGGGCTGCTGTCACCCCTGGTGCC 285
QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSer 100
Db 286 GCGCGCGCGGAGTGCTGGCGCGCTGGCTGGTGGACCGCGCGCGCGCGCGCGCGCT 345
QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAsp 120
Db 346 TTGCTGTCTCGTGCCCTTCGTGGCGCGCTTTCGCTCATCCACCGCGCGCGCGCG 405
QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeu 140
Db 406 TGGATGCTGCTGGGGGCGCGCTCCTCAGCGCGCTGGCGCTGGCGCTGGCTGGCT 465
QY 141 AlaProValTyrIleSerGluLeuAlaTyrProAlaValArgGlyLeuLeuGlySer 160
Db 466 GCCCGCGCTCATCTCCGAAATCGCTACCCAGCAGTCCGCGGCTGCTGGCTGCT 525
QY 161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
Db 526 GTGCAGCTAATGTCGTGTCGGCATCTCTTGGGCTACCTGGCAGGCTGGGTGCTG 585
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200

```

```

586 TGGCGCTGGCTGGCTGTGCTGGGCTGCGTGGCCCTCCCTCATGCTGCTTCTCATGTGC 645
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
Db 646 TTTCATGCCCGAGACCCCGCTTCTGCTGCTGATCAGCACAGGCGCGAGAGCCATGGCC 705
QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProGlyGlyAlaGlu 240
Db 706 GCCCTGGCGGTTCCTGTGGGCTCCAGCAGGAGCTGGAGACCCCTCATGCGGGCTGAG 765
QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleGly 260
Db 766 CAGAGCTTTCACCTGGCCCTGCTGGCGGAGCCCGGCACTACAAACCCCTTCATCATCGGC 825
QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
Db 826 GTCTCCCTGATGGCTTCACGAGCTGTGGGGGTCAACGGCGCTCATGCTTCATGACAG 885
QY 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
Db 886 ACCATCTTTGAAGAGGCCAAGTTCAAGACAGCAGCCTGGCCCTCGTCTGTTGGGTGTC 945
QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeu 320
Db 946 ATCCAGGTGCTTTCACAGCTGTGGCGCTCTCATCATGACAGACGAGGCGGAGGCTG 1005
QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
Db 1006 CTCCTGGTCTTGTACAGGTGTGTCATGTTGTTTACGACAGAGTGCCTTCGGCGCTTCT 1065
QY 341 LysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
Db 1066 AAGCTGACCCAGGGTGGCCCTTGGCACTCTCGACAGTGGCCATCTCGGGCGCTGTCTCT 1125
QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
Db 1126 GCACAGCCTGTGTATGCCAGCGTGGGGCTGGCTGGCTGGCGCGTGGCGAGCATGTGCTC 1185
QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400
Db 1186 TTTCATCGCGCGCTTTCGCGTGGGTGGGGCGCCATCCCTCGCTCTCATGTTCAGAGATC 1245
QY 401 PheProLeuHisValLysGlyValAlaAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
Db 1246 TTCCCTCTGATGTCAAGGGCGTGGCGACAGCATCTCGCTCTCCTCACCACACTGGCTCATG 1305
QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
Db 1306 GCCTTTCGTGACCAAGAGTTTCAGCAGCCTCATGGAGGTCTCTCAGGCCCTATGGAGCC 1365
QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
Db 1366 TTCTGGCTTGGCTCGCTTCTGTCATCTTCATGTCCTTTTTCACCTTCTTCTGTGTCCT 1425
QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 1426 GAACATAAAGGAAGACATCTGGACAAATCAAGCCCATTTTGGGGGCGGA 1476

```

```

RESULT 2
BC019043 2145 bp mRNA linear PRI 19-DEC-2003
LOCUS BC019043
DEFINITION Homo sapiens solute carrier family 2, (facilitated glucose
transporter) member 8, mRNA (cDNA clone MGC:20634 IMAGE:4641145),
complete cds.
ACCESSION BC019043
VERSION BC019043.2 GI:40226554
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2145)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

```

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusha,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Sheets,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettner,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smillius,D.E.,
Schnierch,A., Schein,J.E., Jones,S.J., and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2 (bases 1 to 2145)
Strausberg,R.
Direct Submission
Submitted (07-Dec-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Dec 19, 2003 this sequence version replaced gi:17512129.
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: ATCC/DCTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Iaxic,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Srantripop,S., Thomas,P.J., Touchman,S.W.,
Tsurgoun,C., Voigt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 30 Row: n Column: 3
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 21361448.
Location/Qualifiers
1. .2145
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:20634 IMAGE:4641145"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NIH MGC 20"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1. .2145
/gene="SLC2A8"
/note="synonyms: GLUTX1, GLUT8"
/db_xref="LocusID:29988"
/db_xref="MIM:605245"
46. .1479
/gene="SLC2A8"
/codon_start=1
/product="solute carrier family 2, (facilitated glucose
transporter) member 8"
/protein_id="AAH19043.1"
/db_xref="GI:17512130"
/db_xref="LocusID:29988"
/db_xref="MIM:605245"
/translation="MTPEDEPTOPLLPGPGSAPRRRVFLAAFAAALGPLSPGFAL
GYSPAPLSLORAAAPPRLDDAAASWFAVVTLGAAAGVLLGWLVDRAEKLSLL
CSVPFVAGFAVITAAQDVMLLGRLITGLAGVASLVAPVVISIATPAVRGLGSC
VMAURFLWGSQGWEDPPIGAEOQFHLALLRPGGIYKFIIGVSLMAFQQLSGVNVAM
FYAETIFEEAKFSDSSLASVVGVIQVLTFAAALIMDRAGRLILLVSGVMVFSPT
AFGAYFKLTQGGPGNSHVAISAPVSAQPVDAVSLAWLAVGSMCLFIAGPAVWGPI
PWLMLSEIFPLHVKGVAIGICVLTNMLMAFLVTKEFFSLMEVLRYGAFWLASAFCIF
SVLFTLCVPETKGTLEQITAHEGR"
319. .1467
/gene="SLC2A8"
/note="sugar tr; Region: Sugar (and other) transporter"
/db_xref="CDD:pfam00083"
ORIGIN
Alignment Scores:
Pred. No.: 2,53e-167 Length: 2145
Score: 2457.00 Matches: 477
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-886-954A-1 (1-477) x BC019043 (1-2145)
QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
DB 46 ATGAGCCCGAGGAGCCAGAGAAACCAGCGCGTTCTGGGGCGCTCTCTGGCGGAGCGCG 105
QY 21 ProArgGlyArgArgValPheLeuAlaAalaPheAlaAalaLeuGlyProLeuSerPhe 40
DB 106 CCCCAGCGCGCGCGCTCTCTCGCGCTTCGCGCGCTGCGCGCGCGCGCGCGCGCGCG 165
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
DB 166 GGCTTCGCGCTCGGCTACAGCTCCCGCGCATCCCTAGCTGCAGCGCGCGCGCGCGCG 225
QY 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
DB 226 GCCCCGCGCTGGAGACGCGCGCGCTCTCTGGTTGGGGGCTGCTGACCTGGGTGCC 285
QY 81 AlaAlaGlyGlyValLeuGlyGlyTyrTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
DB 286 GCGCGCGGGGAGTGTCTGGCGCGCTGCTGTGGACCGCGCGCGCGCGCAAGCTGAGCCTC 345
QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
DB 346 TTGCTGTGCTCGGTCGCGCTTGTGGCGCGCTTGTGGCGCTCATCACCGCGCGCGCGCGCG 405
QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
DB 406 TGGATGCTGTGGGGGGCGCGCTCTCACCGCGCTGGCGCTGGGTGTTCCTCCTTAGTG 465
QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
DB 466 GCCCCGCTTACATCTCCGAAATCGCTACCCAGCAGTCCGCGGGGTTGCTCGGCTCCTGT 525
QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGlu 180
DB 526 GTGCAGCTAATGGTCGTCGCGCATCTCTCTGGCGCTACCTGGCAGGCTGGGTCTGTGAG 585
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
DB 586 TGGCGCTTGGCTGTGCTGGGCTGCGTGGCGCGCTTGTGGCGCGCTTGTGGCGCGCTTGT 645
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAla 220
DB 646 TTCATGCCCCGAGACCCGCGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705

FEATURES
source

gene

CDS

```
QY 221 AlaLeuArgPheLeuTyrGlySerGluGlnGlyTyrGluAspProPheGlyAlaGlu 240
Db 706 GCCCTGCGGTTCTCTGTGGGCTCCGAGCAGGGCTGGAGAGACCCCTCCATCGGGCTGAG 765
QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleLeGly 260
Db 766 CAGAGCTTTCACCTGGCCCTGTGCGGACCGCGCATCTACAAGCCCTTCATCATCGGC 825
QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
Db 826 GTCTCCCTGATGGCTTCCAGCAGCTGTGCGGGGTCAACGCCGTGATGTTCTATGACAG 885
QY 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
Db 886 ACCATCTTTGAAGAGCCGACAGTTCAAGACACAGCTGCGCTCGGTGCTGTGGTGTTC 945
QY 301 IleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgGlyLeu 320
Db 946 ATCCAGGTGCTGTTACAGCTGTGGCGGCTCTCATCGACAGAGCGCGGAGGCTG 1005
QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
Db 1006 CTCCTGGTCTTGTGAGTGTGTGATGGTGTTCAGCAGAGTGCCCTTCGGCGCCCTACTTC 1065
QY 341 LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
Db 1066 AAGCTGACCCAGGGTGGCCCTGGCACTCTCGCAGCTGGCCATCTCGCGGCTGTCTCT 1125
QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTyrPheLeuAlaValGlySerMetCysLeu 380
Db 1126 GCACAGCCTGTTGATCCAGCGTGGGGCTGGCGTGGCTGGCGCAGCATGTGCTCTC 1185
QY 381 PheIleAlaGlyPheAlaValGlyTyrGlyProIleProTyrLeuLeuMetSerGluIle 400
Db 1186 TTCATCGCGGCTTTCGCGTGGGCTGGGGGCCATCCCTGGCTCCTCATGTCAGAGATC 1245
QY 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTyrLeuMet 420
Db 1246 TTCCTCTGCATGTCAAGGGCTGGGCGCAGGCACTGCGTCTCACCACTGGCTCATG 1305
QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
Db 1306 GCCTTTCCTGTCACCAAGGAGTTCAAGCGCTCATGAGGCTCCTCAGGCGCTATGGAGCC 1365
QY 441 PheTyrLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
Db 1366 TTCCTGCTTGCCTTCGCTTTCGATCTTCAGTGTCTTTTTCACCTTGTCTGTGTCCT 1425
QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 1426 GAAACTAAAGGAAAGACTCTGGAAACAAATCACAGCCCATTTTGAGGGGCGA 1476

RESULT 3
HSA245937
LOCUS HSA245937 1873 bp mRNA linear PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA for glucose transporter (GLUTX1 gene).
ACCESSION AJ245937
VERSION AJ245937.1 GI:7018305
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Ibberson,M., Uldry,M. and Thorens,B.
TITLE GLUTX1, a novel mammalian glucose transporter expressed in the central nervous system and insulin-sensitive tissues
JOURNAL J. Biol. Chem. 275 (7), 4607-4612 (2000)
MEDLINE 20138191
PUBMED 10671487
REFERENCE 2 (bases 1 to 1873)
AUTHORS Ibberson,M.R.
TITLE Direct Submission
```

JOURNAL

Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne, CH-1005, SWITZERLAND

FEATURES

source

Location/Qualifiers

1..1873
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1873
/gene="GLUTX1"
4..1437
/gene="GLUTX1"
/codon_start=1
/product="glucose transporter"
/protein_id="CAB75702.1"
/db_xref="GI:7018306"
/db_xref="GOA:Q9NY64"
/translation="MTPEDEPTQLLPGPPGGSAPRRRVFLAAFAAALGPLSFGFAL
GYSSPAISLQRAAPPAPRLDDAAASWFGAVVTLGAAAGVGLGWLVDRAGRKLSLL
CSVPEVAGFAVITAQDYVMLLGGRLLTGLAGVASLAVPVVISIAYPAVRGLGSC
VQLMVVGIILAYLAWGLEWELAVLGCVPSPMLLLCMFPEPTEPRELLTQHRQEA
MAALRVFWSQGWEDPPIGAESPHLALLKQPGYIKFIIIGVSLMAFQQLSGVNAV
FYAETIFBEAKFKDSSLASVVGVIQVLFATAALIMDRAGRLLLVSGVVMVST
AFGAYFKLTQGGPGNSHVAISAPVASGLAVGNMCLFTAGFAVGRGPI
PWLLMSEIPLHLVKGVATGICVLTNMLMAFLVTKFSSILMEVLRYPGAPWLASAF
SVLFTLCVPEIKGKLTLEQITAHFGR"

ORIGIN

Alignment Scores:

Pred. No.: 9.55e-167 Length: 1873
Score: 2448.00 Matches: 475
Percent Similarity: 99.79% Conservative: 1
Best Local Similarity: 99.58% Mismatches: 1
Query Match: 99.63% Indels: 0
DB: 9 Gaps: 0

US-09-886-954A-1 (1-477) x HSA245937 (1-1873)

```
QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProGlyGlySerAla 20
Db 4 ATGAGCCCGAGAGACCCAGAGAAACCCAGCGCTTCCTGGGGCCCTCCTGCGCGACGCG 63
QY 21 ProArgGlyArgValPheLeuAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
Db 64 CCCCAGGCGCGCGGCTTCTCTCGCGCTTCGCGCTGCCCTGGGCCACTCAGCTTC 123
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaProPro 60
Db 124 GGCTTCGGCTCGGCTACAGCTCCCGGCCATCCCTAGCTCAGCGCGCGCGCCCG 183
QY 61 AlaProArgLeuAspAlaAlaAlaSerTyrPheGlyAlaValValThrLeuGlyAla 80
Db 184 GCGCGCGCTCGAGCAGCGCGCGCTCTCGTTCGGGGCTGTGTGACCTGGGTGCC 243
QY 81 AlaAlaGlyGlyValLeuGlyGlyTyrLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db 244 GCGCGCGGGAGTGTCTGGCGGCTGGCTGGTGACCGCGCGCGCGCAAGCTGAGCTTC 303
QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db 304 TTGCTGTGCTCGGCTCGCTTCGTGGCGGCTTTGCCGTTCATCAGCGCGCGCGAGCAGTG 363
QY 121 TrrMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db 364 TGGATGCTCTCGGGGGCGCGCTCTCTACCGGCTGGCTCGGGTGTGCTTCCTTAGTG 423
QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
Db 424 GCGCGGCTCTACATCTCCGAAATCGCTACCCAGCAGTCCGGGGGTGCTCGGCTCTGT 483
QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrrValLeuGlu 180
```



```
Db 390 TGGATGCTGCTCGGAGCGCGCTCTCTACCGGCTAGCTCGGAGTGCCTCATTAGT 449
QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuGlySerCys 160
DB 450 GCACCGGCTATATCTCGAATTCGCTACCCAGCGCTTCGAGGACTGCTCGGCTCCTGT 509
QY 161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGlu 180
DB 510 GTGAGCTGATGGTGTCTACTGGCATCTCTTGGCTAGCTGGCAGGCTGGGTCTAGNA 569
QY 181 TtrArgTrrLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
DB 570 TGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 629
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
DB 630 TACATGCCCGAGACCCACGCTTCTCTGAGCGAACACACCATACAGAGGCCATGGCT 689
QY 221 AlaLeuArgPheLeuTrrGlySerGluGlnGlyTrrGluAspProProIleGlyAlaGlu 240
DB 690 GCCCTGCGCTTCTCTGGGCTCTGAGGAGGCTGGAGAGCCCTGTTGGGCTGAG 749
QY 241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIysProPheIleIle 259
DB 750 CACCAGGGCTTCCAGCTGGCTATGCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 809
QY 260 GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAlaAlaValMetPheTyrAla 279
DB 810 GGCATTGCTCATGCTTCTCAGCAGCTGCTCAGGAGGCTGAGGAGGCTGAGGAGGCTG 869
QY 280 GluThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGly 299
DB 870 AACACCATCTTGGAGGCGCAAGTTCAGGACAGCAGCTGCTGCTGCTGCTGCTGCTG 929
QY 300 ValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArg 319
DB 930 ATCATCCAGGTTCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 989
QY 320 LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr 339
DB 990 CTGCTTCTGCGCTTCTGCGGTGTGATCGGTTCAGATGAGGCGCTTCTGCTGCTGCT 1049
QY 340 PheLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProVal 359
DB 1050 TTCAAACTGACCCAGAGTGCCGCCAGCACTCTCCATGTAGGCTCCTGCTGCTGCTG 1109
QY 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTrrLeuAlaValGlySerMetCys 379
DB 1110 TCCGAGAGCTGCTGATGTTTCCCTGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 1169
QY 380 LeuPheIleAlaGlyPheAlaValGlyTrrGlyProIleProTrrLeuLeuMetSerGlu 399
DB 1170 CTCCTTCATCGCTGGTTCATGATGCTGGGACCCATCCCTGCTGCTGCTGCTGCTGCTG 1229
QY 400 IlePheProLeuHisValIysGlyValAlaThrGlyIleCysValLeuThrAsnTrrLeu 419
DB 1230 ATCTTCCCTCTGCATCATCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1289
QY 420 MetAlaPheLeuValThrIysGluPheSerSerLeuMetGluValLeuArgProTrrGly 439
DB 1290 ATGGGCTTCTGCTGACCAAGAGTTTACAGCATCATGAGATCCTCAGACCCCTACGCG 1349
QY 440 AlaPheTrrLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
DB 1350 GCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1409
QY 460 ProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
DB 1410 CCTGAGACTAAAGGCGAGGACTCTGGAACAAATCACAGCCCAATTCGAGGGGCGG 1463
RESULT 7
RNO245935
LOCUS
```

```
DEFINITION Rattus norvegicus mRNA for glucose transporter (GLUTX1 gene).
ACCESSION AJ245935
VERSION AJ245935.1 GI:7018604
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE 1
AUTHORS Ibberson,M., Uldry,M. and Thorens,B.
TITLE GLUTX1, a novel mammalian glucose transporter expressed in the
        central nervous system and insulin-sensitive tissues
JOURNAL J. Biol. Chem. 275 (7), 4607-4612 (2000)
MEDLINE 20138191
PUBMED 10671497
REFERENCE 2 (bases 1 to 2087)
AUTHORS Ibberson,M.R.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
        and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
        CH-1005, SWITZERLAND
FEATURES             Location/Qualifiers
     source           1..2087
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
     gene             1..2087
                     /gene="GLUTX1"
                     /length=2087
     CDS               30..1466
                     /gene="GLUTX1"
                     /codon_start=1
                     /product="glucose transporter"
                     /protein_id="CAB57529.1"
                     /db_xref="GI:7018605"
                     /db_xref="SWISS-PROT:Q9JUZ1"
                     /translation="MSPEDPQETPTLLRSPGARPGRRVFLATFAAALGPLSGFAL
                     GYSPALPSLRRTAPPALRLGDTAASWFGAVVTLGAAAGVLGWLMDRAGRKLKLL
                     CTVPVTGPAVITAARDVWMLGGRLTGLACGASLVAPVVISIAYPAVGLGSC
                     VOLMVVTGILLAYVAGVLEWLAVALGCVPTLLMLLCCMPETPRLLTTHQHQIEA
                     MAALFLMGSEEGWEEPPVGAHQFQAMLRPGVHKPLIIGICLMVYQQLSGVNAI
                     MFYANTIPFEAKQDSLSAVTVGIQVLTAVALIMDRAGRLKLLALSIVMVFMS
                     SAFGTYPKLQSGSPNSHVLGLYISAEPADVHLGLAWLAVGSMCLFATGFAVGWGP
                     IPWLMSRIPLHLKGVATGCVLTNWFMAFLVTKEFNSIMEILRPYGAFWLTAACFI
                     LSVLTFLTFVETKGRILEQITAHEG"
ORIGIN
Alignment Scores:   3.51e-146   Length:      2087
Pred. No.:         2162.50   Matches:    408
Score:             92.2%     Conservative: 33
Percent Similarity: 85.3%    Mismatches:  36
Best Local Similarity: 88.01% Indels:      1
Query Match:       10       Gaps:      1
DB:
US-09-886-954A-1 (1-477) x RNO245935 (1-2087)
QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
DB 30 ATGTCGCCCGAGGAGCCCGACGAAACACAGCGCGCTGTTGCGGTGCGCGCGCCAGGCT 89
QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
DB 90 CCCGGCGCGCGCGGCTTCTCTTCCACCTTCGCGCGCGCGCTGCGCGCGCTCAGCTC 149
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
DB 150 GGCTTCGCGCTGGCTACAGCTCCCGCCCATCCCGAGCTCGCGCGCGCACCCGCT 209
QY 61 AlaProArgLeuAspAlaAlaSerTrrPheGlyAlaValValThrLeuGlyAla 80
DB 210 GCCCTACCGCTCGGAGACACTGCGGCGCTCTCTGTTCCGGCGCGCTGCTGCGCGCT 269
```

```
Qy 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db 270 GCGCAGGGGCGTGTGGGCGGTGGCTCTCTGACCGCTGCGAGCGCAAGCTGAGGCTC 329
Qy 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db 330 CTGCTCTGCACCGTGGCTGCTGTCGACCGGTTTGTCTCATCACCGCGGCCGGGAGTGTG 389
Qy 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaAlaSerLeuVal 140
Db 390 TGGATGCTGCTCGAGCGCGCTCTCTCACCGCGCTAGCCTGCGAGTGCCTCATTAGTG 449
Qy 141 AlaProValTyrlleSerGluIleAlaTyrlProAlaValArgGlyLeuLeuGlySerCys 160
Db 450 GCACCGGCTCATATCTCGGAATCGCTACCGCGCTTCGAGAGCTGCTCGGCTCCTGT 509
Qy 161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrlLeuAlaGlyTrpValLeuGlu 180
Db 510 GTGAGCTGATGTTGTCTACCTGGCATCTCTTGGCTTACGTGGCAGGCTGGTCTTAGAA 569
Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 570 TGGCGCTGGTGGCTGCTGGGCTGTGTGCGGCCCTCACCTCATGCTGCTCATGTGC 629
Qy 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
Db 630 TACATGCCGAGACCCACCGCTTCTCTGACGCAACACCATGATCCAGGAGCGCATGGCT 689
Qy 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProlleGlyAlaGlu 240
Db 690 GCCCTGGCTTCTGTGGGCTCTGAGAGGGCTGGGAAGAGCCCTGTGTGGGGCTGAG 749
Qy 241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrlLysProPheIle 259
Db 750 CACCAGGCTTCCAGCTGGCTATGCTGAGGCGCCTGGTGTCCACAAGCCCTCATCATC 809
Qy 260 GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAanAlaValMetPheTyrlAla 279
Db 810 GGCATTTGGCTCATGGTCTTCCAGCAGCTGTCCAGGGGTCAACGCGCATCATGTTCTATGCC 869
Qy 280 GluThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGly 299
Db 870 AACACCATCTTGGAGAGCCCAAGTTCAAGACAGCAGCGCTGGCTGCTGCTGCTGGGC 929
Qy 300 ValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgArg 319
Db 930 ATCATCCAGGTTCTGTTCACTGCTGTGGCGCCCTCATCATGGACAGAGCGCGGAA 989
Qy 320 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrl 339
Db 990 CTGCTTTCGGCTTGTGGGTTGTATCATGTTTTCACATGAGCGGCTTTGGTACCTAC 1049
Qy 340 PheLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProVal 359
Db 1050 TTCAACTGACCCAGAGTGGCCCCAGCACTCTCTCCATGTAGGCGCTTCTGGTCCCATC 1109
Qy 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCys 379
Db 1110 TCCGACAGCGCTGCTGATGTTCACTGGGGTGGCTGGCTGGCTGTAGGCAGCATGTGC 1169
Qy 380 LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGlu 399
Db 1170 CTCCTTCATCGCTGTTTTCAGTAGTGGGAGCCCATCCCTCGCTCCTCATGTACAG 1229
Qy 400 IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeu 419
Db 1230 ATCTTCCCTCTGCACATCAAGGGTGTGGCTACCGCGCTGTGTGCTTCCACCACTGGTTC 1289
Qy 420 MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrlGly 439
Db 1290 ATGGCTTCTGTTGACCAAGAGTTTAAACAGCATCATGGAGATCCTCAGACCCCTACGCC 1349
```

```
Qy 440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
Db 1350 GCCTTCTGGCTACCGCTGCTCTCTGTATCTCAGCGTCTCTTTCACGCTACACTTGTGTC 1409
Qy 460 ProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 1410 CCTGAGACTAAGCAGGACTCTGGAACAATCACACCCATTCGAGGACGG 1463

RESULT 8
MMU17802 1490 bp mRNA linear ROD 13-MAY-2000
LOCUS Mus musculus mRNA for glucose transporter 8 (GLUT8 gene).
DEFINITION Y17802
ACCESSION Y17802
VERSION Y17802.1 GI:7688219
KEYWORDS glucose transporter 8; GLUT8 gene.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1
AUTHORS Doege,H., Schurmann,A., Bahrenberg,G., Brauers,A. and Joost,H.G.
TITLE GLUT8, a novel member of the sugar transport facilitator family
with glucose transport activity
J. Biol. Chem. 275 (21), 16275-16280 (2000)
JOURNAL 20283667
MEDLINE 10821868
PUBMED 2 (bases 1 to 1490)
REFERENCE Joost,H.G.
AUTHORS Direct Submission
TITLE Submitted (16-JUL-1998) H. Joost, Institute of Pharmacology and
Toxicology, Technical University Aachen, Wendlingweg 2, D-52057
Aachen, FRG
FEATURES
source 1. .1490
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="testis"
1. .1490
/genes="GLUT8"
26. .1459
/genes="GLUT8"
/codon_start=1
/product="glucose transporter 8"
/protein_id="CAB89815.1"
/db_xref="GI:7688220"
/db_xref="GOA:Q9JIF3"
/db_xref="SWISS-PROT:Q9JIF3"
/tranlation="MSPEDPQETQPLLRPPEARTPRGRRVFLASFAALGPLSPGFAL
GYSSPAISLRRTAPALRLGDNAASFVAVTLGAAAGILGWLDRGRKLSLLL
CTVFTGFAVITTAARDVWMLLGGRLTLGLACGVASLVAPVISEIAYPAVRGLLGC
VQLVMTGILLAYVAGVLEWRWLAVLGVPTLMGLMCMYMPETPRFLITQHQYQA
MAALRLFWSEGEWEPVGAHQGLALLRRPGYKPLIIGTSLMVFOQLSGVNAI
MFVANSIPEAKKDSLSASTVGIQVLFVAVALTMDRAGRLLIALSGVINVFQM
SAGTGFKLQSLPSNSSSHVGLPIAEAPVDVQGLAVAGSMCLFIAGFVAGGPI
PWLIMSEIFPLHVKGVATGICVLINPMAFLVTKFSSVMEMLRPHYCAFWLTAAFCAL
SVLFTLTUVVPETKGRTLEQVTAHFEGR"
ORIGIN
Alignment Scores:
Pred. No.: 4,19e-146 Length: 1490
Score: 2159.00 Matches: 412
Percent Similarity: 92.47% Conservative: 30
Best Local Similarity: 86.19% Mismatches: 34
Query Match: 87.87% Indels: 2
DB: 10 Gaps: 2
US-09-886-954A-1 (1-477) x MMU17802 (1-1490)
```

```
Qy 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProGlyGlySerAla 20
Db 26 ATGTCTCCGAGGACCCCGAGGAGACGACCGCTATTGCGGCCACCGAAGCAGGACT 85
```

QY 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 DB 86 CCGCGGCGCGGGTCTTCCTCGCTTCCTTCGCGCGCGCTCTGGACCCCTCAGCTTC 145
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaPro 60
 DB 146 GCGTTTCGCGCTCGGCTACAGCTCCCGCGCATCCCGCGCTCGCGCGCACGCCCGC 205
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 DB 206 GCGCTCGCGCTCGGAGCAATCGCGCTTCCTGGTTCGGGCGCGCTGTACCGCTGGCGCT 265
 QY 81 AlaAlaGlyGlyValLeuGlyTyrTrpLeuValAspArgAlaGlyArgLeuSerLeu 100
 DB 266 GCTGACGGGGCATACTGGCGGCTGGCTCTCGACCGCTGACGGCGCAAGCTGAGCCTC 325
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 DB 326 TTGCTCTGCACCGTCCCTTCGTGACTGGCTTGTCTGTCATCACCAGCGCGCGGATGTG 385
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 DB 386 TGGATCTGCTCGGAGCGCGCTCTCACCAGCGCTTGGCTGCGAGTGCCTCCTCAGTAGT 445
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 DB 446 GCACCGGTTTACATCTCGGAATCGCTACCCAGCTGTCGAGGACTGCTCGGCTCCTGT 505
 QY 161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGlu 180
 DB 506 GTGCAGCTGATGGTTGTCACTGGCATCTCTCGGCTATGTGGCAGGCTGGGTCTCAGAG 565
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 DB 566 TGGCGCTGGCTGGCGCTGCTGGCTGTGTGCCCCCACCCTCATGCTGCTCATGTGC 625
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 DB 626 TACATGCCGAAACCCACGTTTCTCTCACTCAACACCAGTACACAGAGGCCATGGCT 685
 QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGlu 240
 DB 686 GCTTTGCGCTTCCTGTGGGGCTCTGAGGAGGGCTGGGAAGAGCCCGCTGTGGGGCTGAG 745
 QY 241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIle 259
 DB 746 CACCAGGGCTTCCAGCTGGCGCTGCTGAGGGCCCTGCACCTACACCCCTCATCATC 805
 QY 260 GlyValSerLeuMetAlaPheGlnLeuSerGlyValAsnAlaValMetPheTyrAla 279
 DB 806 GGCATTTCCCTCATGGTCTTCCAGCAGCTGTCAAGGGTCAATGCTATCATGTTCTATGCC 865
 QY 280 GluThrIlePheGluGluAlaIlePheIleYsPheSerSerLeuAlaSerValValGly 299
 DB 866 AACAGCATCTTCGAGGAGGCCAAGTTCAGGACAGCAGCCCTGGCGCTCGGTCACTGTGGGC 925
 QY 300 ValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArg 319
 DB 926 ATAATCCAGGTCCTGTTCACTGCTGTGGCGCCCTCATCGGACAGAGCGAGCGAAGG 985
 QY 320 LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr 339
 DB 986 CTGCTCTCGGCTTGTGCGGTGTGATCATGTGTGTTCAGTAGAGTGGCTTTGGTACCTAC 1045
 QY 340 PheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProVal 359
 DB 1046 TTCAACTGACCGAGAGCTCCCGAGCACTCTCCACGAGGCGCTG---GTGCGCCATC 1102
 QY 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCys 379
 DB 1103 CGCGCGGAGCGCTGTGGATGTCCAAGTGGACTGCGCTGGCTGGCTGAGGAGCATGTGC 1162
 QY 380 LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGlu 399

DB 1163 CTCCTCATGTGGCTTTGGGTGGGAGCCATCCCTGGCTCTCATGTACAGAG 1222
 QY 400 IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeu 419
 DB 1223 ATCTTCCCTCTGCATGTCAAGGGTGTGGCTACCGGCATCTGTGTCTCACCACCTGGTTC 1282
 QY 420 MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGly 439
 DB 1283 ATGGCTTTTCTAGTGACCAAGAGTTTCAGCAGCGTCATGGAGATGCTCAGACCTACCGT 1342
 QY 440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
 DB 1343 GCCTTCTGGCTACCGCTGCTCTTGGCGCTCTCAGTGTCTTATTACACTGACCTGTC 1402
 QY 460 ProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 DB 1403 CCTGAGACTAAAGGCGAGGACTCTGGAAACAAGTCAAGCCCATTTCCAGGGACGA 1456
 RESULT 9
 AF232061 1843 bp mRNA linear ROD 23-JUN-2000
 LOCUS Mus musculus glucose transporter GLUT8 mRNA, complete cds.
 DEFINITION AF232061
 ACCESSION AF232061
 VERSION AF232061.1 GI:8671757
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1843)
 AUTHORS Carayannopoulos M.O., Chi, M.M., Cui, Y., Pingsterhaus, J.M.,
 McKnight, R.A., Mueckler, M., Devaskar, S.U. and Moley, K.H.
 TITLE Glut8 is a glucose transporter responsible for insulin-stimulated
 glucose uptake in the blastocyst
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (13), 7313-7318 (2000)
 MEDLINE 20319023
 PUBMED 10860996
 REFERENCE 2 (bases 1 to 1843)
 AUTHORS Moley, K.H., Carayannopoulos, M.O. and Cui, Y.
 DIRECT SUBMISSION
 TITLE Submitted (07-FEB-2000) OB/Gyn, Washington University, 4566 Scott
 JOURNAL Ave, St. Louis, MO 63110, USA
 FEATURES
 source
 1..1843
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129"
 /db_xref="taxon:10090"
 1..-48
 49..1482
 /codon_start=1
 /product="glucose transporter GLUT8"
 /protein_id="AAF78366.1"
 /db_xref="GI:8671758"
 /translation="MSPEDQETQPLLRPPPEARTPGRRRVFLASFAALGFLSPGFL
 GYSPALPSLRRTAPPALRLGDNAAMFEGAVVTLGAAGGILGGLDGRKLSLL
 CTVPPTGPAVITAARDVWMLLGRILITGLACVASLVAIPYISEIAYPAVRGLGSC
 VQLMVVTGILLAYVAGVLEWRMLAVLGVCPPTLMLLMCMTPETPRFLITQHGYQEA
 MAALRFLWGSEEGEPPVGAHGQFQALLRRPGIYKPLIIIGLSLMVFOQLSGVNAI
 MFANSIFEEAKFKDSSLASVTWGIIQVLTAAALIMDRALLLALSGVIMVFSM
 SAFQTYFKLTQSLPSNSSHVGLVPIAAEPDVQVGLAWLAVGSMCLFTAGFVGWGP
 PWLLMSRIFLHVKGATGICVLTFWPMFLVTKFSSVMEMLRPYGAFLWLTAAFCAL
 SVLPTLVVPETKGRTLEQVTAHFEGR"
 1483..1843
 3'UTR
 ORIGIN
 Alignment Scores:
 Pred. No.: 8.89e-146 Length: 1843
 Score: 2156.00 Matches: 411
 Percent Similarity: 92.47% Conservative: 31
 Best Local Similarity: 85.98% Mismatches: 34
 Query Match: 87.75% Indels: 2

DB: 10 Gaps: 2

US-09-886-954A-1 (1-477) x AF232061 (1-1843)

Qy 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db :
49 ATGAGTCCCGAGGACCCCAGGAGACTCAGCGCTATTGGCGGCACC CGGAAGCAGGACT 108
Qy 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
Db :
109 CCCGCGCGCGCGGTCTTCCTCGCTTCCTTCGCCGCGCTCTGGGACCCCTCAGCTTC 168
Qy 41 GlyPheAlaLeuGlyTyrrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
Db GGCTTCGCGTTCGCTACAGTCCCCCGCATCCACAGCCTGGCGGCACCGCACCCCGC 228
Qy 61 AlaProArgLeuaspAspAlaAlaAlaserTrpPheGlyAlaValValThrLeuGlyAla 80
Db :
229 GCCCTGCGCTTCGAGACAATGCGGCTCTCTGGTTCGGGCGCGTGFAGACCTCGGGCGCT 288
Qy 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db :
289 GCTGCGAGGGGCATACCTGGGCGGTGCTCTCGACCGTTTCAGGGCGCAAGCTCAGCTC 348
Qy 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db :
349 TTGCTCTGCACCGTGCGCTTCGTGACTGGCTTTGCTGTCATCACCGCGGCCCGGATGTG 408
Qy 121 TrpMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaAlaSerLeuVal 140
Db :
409 TGGATGCTGCTCGAGGCGCGCTCTCACCGGCTTCGCTCGGAGTCGCTCCTCCTAGTG 468
Qy 141 AlaProValTyrrIleSerGluIleAlaTyrrProAlaValArgGlyLeuLeuGlySerCys 160
Db GCACCGGTTTACATCTCGAAATCGCCTACCCAGCTGTCGAGGACTGCTCGGCTCTCTGT 528
Qy 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrrLeuAlaGlyTrpValLeuGlu 180
Db GTGAGCTGATGTTGTCACTGGCATCTCTCGGCTATGTGCGAGCTGGGCTCGCTTAGAG 588
Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db TGGCGCTGGCTGGCGCGTGTGGGCTGTGCCCCCACCCTCATGCTGCTGCTCATGTGC 648
Qy 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
Db TACATGCCGAGACCCACGTTTTCTCTCCTCCTCACTCAACACCACTACCAAGGAGGCATGGCT 708
Qy 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
Db GCCTTGGCTTCCTGTGGGCTCTGAGAGGGCTGGGAAGGCCCCCTGTTGGGCTGAG 768
Qy 241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrrIysProPheIleIle 259
Db CACCAGGCTTCCAGCTGGCCGCTGCTGAGCGGCCCTGGCATCTACAAAGCCCCCTCATCATC 828
Qy 260 GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaAlaMetPheTyrAla 279
Db GGCAATTCCCTCATGCTTCCAGCAGCTGTACGGGTCAATGCTATCAITGTTCTATGCC 888
Qy 280 GluThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGly 299
Db AACAGCATCTTCGAGGAGGCCAAGTTCAGGACAGCAGCTGGGCTCGGTCACCTGTGGGC 948
Qy 300 ValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArg 319
Db ATAATCAGGTCTGTTTCTGCTGTGGCGCCCTCATCTATGACACAGCAGCGCGCAAGG 1008
Qy 320 LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr 339
Db CTGCTCTCGGCTTGTGGGTGTGATCATGGTTCAGTATGAGTGTGCTTTGGTACCTAC 1068
Qy 340 PheLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProVal 359

US-09-886-954A-1 (1-477) x AX076671 (1-2072)

```
QY      1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlySerAla  20
Db      21 ATGTCTCCCGAGGACCCACGAGACGACGCGCTATTGCGGCCACCGGAACCCAGGACT  80
QY      21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe  40
Db      81 CCCCGCGGCGCGGGTCTTCTCGCTTCCTTCGCGCGCGCTCTGGACCCCTCAACTTC  140
QY      41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaProPro  60
Db      141 GCGTTTCGCGCTCGGCTACAGCTCCCGCGCATCCCGCAGCTCGCGCGCACCCGCCCG  200
QY      61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla  80
Db      201 GCGCTCGCGCTCGGAGCAATCGGCGCTCTCGTTTCGGGCGCGCTGATGACCTGGCGCT  260
QY      81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlySerLeu  100
Db      261 GCTCAGGGGGCATACTGGCGGCTGGCTCTCTGGACCGTTTCAGGGCGCAAGCTGAGCCTC  320
QY      101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal  120
Db      321 TTGCTCTCGACCGCTCCCTTCGTGACTGCTTGTCTTCATCACCAGCGCGCGGATGTG  380
QY      121 TrpMetLeuLeuGlyGlyValGluLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal  140
Db      381 TGGATGCTGCTCGGAGGCGGCTCTCTACCGGCGCTTGCCTGGGAGTGCCTCACTAGTG  440
QY      141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys  160
Db      441 GCACCGGTTTACATCTCGAAATCGCTACCCAGCTGTCGAGGACTGCTCGGCTCTGT  500
QY      161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu  180
Db      501 GTGAGCTGATGTTGTCACTGGCATCTCTCTGGCTATGTCGCGGCTGTCGAGGCTGAG  560
QY      181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys  200
Db      561 TGGCGCTGGCTGGCGCTGCTGGGCTGTGTGCCCCCACCCTCATGCTGCTGCTCAATGTC  620
QY      201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla  220
Db      621 TACATGCCGAGACCCACGCTTTCTCTCACTCAACACCACTAGTACGAGGCCCATGGCT  680
QY      221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu  240
Db      681 GCGTTGCGCTTCCTGTGGGCTCTGAGGAGGCTGCGGAGAGCCCGCTGTTGGGCTGAG  740
QY      241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIle  259
Db      741 CACAGGGCTTCCAGCTGGCCCTGCTGAGGCGCCCTGGCACTACAGCCCCCTCATATC  800
QY      260 GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla  279
Db      801 GGCATTTCCTCATGCTCTCCAGCAGCTGTCCAGGGTCAATGCTATCATGTTCTATGCC  860
QY      280 GluThrIlePheGluGluAlaIysPheIysAspSerSerLeuAlaSerValValGly  299
Db      861 AACAGCATCTTCGAGAGGGCCAAAGTTCAAGGACAGACGCGCTGCGCTCAGCTGGGCG  920
QY      300 ValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgArg  319
Db      921 ATAATCCAGGTCCTGTTCACTGCTGTGGGCGCCCTCATCATGACAGAGCGGCGCAAG  980
QY      320 LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr  339
Db      981 CTGCTCCTGGCCCTTCGCGGTGTGATCATGTGTGTTCAGTATGAGTGCCTTTGGTACCTAC  1040
QY      340 PheLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProVal  359
Db      1041 TTCAAACAGACCCAGAGCCCTCCCGACCACTCCTCCACGAGGCGCTG---GTGCCCCATC  1097
```

```
QY      360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCys  379
Db      1098 GCGGCGGAGCGCTGTGGATGTCCAGTGGGACTGGGCTGGCTGGCTGTAGGACGATGTC  1157
QY      380 LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGlu  399
Db      1158 CTCCTTCATGCTGGCTTTGCGGTGGGCTGGGACCCATCCCTGGCTCTCATGTGTCAGAG  1217
QY      400 IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeu  419
Db      1218 ATCTTCCTCTGCATGTCAAGGGTGGCTACCGGCATCTGTGTCTCCACCACTGGTTC  1277
QY      420 MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTrpGly  439
Db      1278 ATGGCTTTCTAGTGACCAAGAGTTCAACAGCGTCATGGAGATGCTCAGACCCCTACCGT  1337
QY      440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal  459
Db      1338 GCCTTCGCTCACCGCTGCCTTCTGCGCTCTCAGTGTCTTATTACACTGACCGCTGTGC  1397
QY      460 ProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg  477
Db      1398 CCTGAGACTAAAGGCGAGGACTCTGGAAACAAGTCACAGCCCATTTCCAGGAGCA  1451

RESULT 11
LOCUS   MMU245936                      2072 bp      mRNA      linear      ROD 18-FEB-2000
DEFINITION Mus musculus mRNA for glucose transporter (GLUTX1 gene).
ACCESSION AJ245936
VERSION   AJ245936.1 GI:7018606
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS  Ibbersson, M., Uldry, M. and Thorens, B.
TITLE    GLUTX1, a novel mammalian glucose transporter expressed in the
JOURNAL  J. Biol. Chem. 275 (7), 4607-4612 (2000)
MEDLINE  20138191
PUBMED   10671487
REFERENCE 2 (bases 1 to 2072)
AUTHORS  Ibbersson, M.R.
TITLE    Direct Submission
JOURNAL  Submitted (02-SEP-1999) Ibbersson M.R., Institute of Pharmacology
and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
CH-1005, SWITZERLAND
FEATURES             Location/Qualifiers
     source           1..2072
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
     gene             1..2072
                     /gene="GLUTX1"
     CDS              21..1454
                     /gene="GLUTX1"
                     /codon_start=1
                     /product="glucose transporter"
                     /protein_id="CAB75719.1"
                     /db_xref="GI:7018607"
                     /db_xref="GOA:Q9JIF3"
                     /db_xref="SWISS-PROT:Q9JIF3"
                     /translation="MSFDPQETQFLRPPPTPRGRVRFLASFAALGRLNFGFAL
GYSPALPSRRRTAPPALRUDNAAWFGAVVTIGAAAGILGGMLLDRSGRKLULL
CTVPFTVGTGVTIATARDVMMLLGRLITGLACVASLVPVYIETIAYPAVRGLGSC
VOLMVVTGILLAYVAGWLEWRWLAVLGVCPPTLMLLMCMYMPETPRLLTHQYQEA
MAALRWLWSEGEPEEPVGAEHQGFQALLRRPGIKYKPLIIGISLMYFQQLSGVNAI
MFYANSIFERAKPKDSSIASVTVGIIQVLTFAVALIMDRAGRRLLALSGVIMVFSM
SAFTYKLTQSLPSNSHVGVLPIAEPPDVQVGLAWLAVGSCMLFIAGFAVAGWPI
FWLWMSEIFLPHVKGATGICVLTNWFMFLVTKFNFSVMENMLRPGYGFALTAAPCAL
SVLPTLTIVVPETKGRITLEQVTAHFEGR"
```

ORIGIN

Alignment Scores:

Pred. No.: 2,75e-145 Length: 2072
 Score: 2150.00 Matches: 409
 Percent Similarity: 92.47% Conservative: 33
 Best Local Similarity: 85.56% Mismatches: 34
 Query Match: 87.51% Indels: 2
 Ds: 10 Gaps: 2

US-09-886-954A-1 (1-477) x MMU245936 (1-2072)

QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 Db 21 ATGTCTCCGAGGACCCCGAGGAGCGAGCGCGCTATTTCGGGCCACCGGAACCCAGGACT 80
 QY 21 ProArgGlyArgArgValPheLeuAlaAalaPheAlaAlaLeuGlyProLeuSerPhe 40
 Db 81 CCCCAGCGCGCGCGGCTCTTCCTCGCTTCTTCGCGCGCGCTCTTCGGACCCCTCAACTTC 140
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 Db 141 GCCTTCGGCGCTCGGTACAGCTCCCGCGCATCCCGAGCTCGCGCGCACCGCACCCCGG 200
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 Db 201 GCCCTGCGCCTCGGAGACAAATCGCGCCTCTCGTTTCGGGGCGCTCGTACCTTCGGCGCT 260
 QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlySerLeu 100
 Db 261 GCTCAGGCGGCGATACGCGCGCGCTCGGTCTCGGACCGTTTCAGGCGCGCAAGCTGAGCCTC 320
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 Db 321 TTGCTCTGCACGTCGCTTCGTGACTGCTTCGTCTCATCACCGCGCGCGCGGATGTG 380
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaAlaSerLeuVal 140
 Db 381 TGGATGCTGCTCGGAGCGCGCTCTCACCGGCTTGCTCGCGAGTCGCTCACTAGTG 440
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 Db 441 GCACCGGTTTACATTCGGAATCCCTACCCAGCTGTCCGAGGACTGCTCGGCTCTCTGT 500
 QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 Db 501 GTGACAGTATGGTTGTCTACTGGCATCTCTCGGCTATGTGCGAGGCTGGTCTCTAGAG 560
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 Db 561 TGGCGCTGGCTGGCGCTGTGGCTGTGTGCCCCCACCTCATGCTGCTGCTCATGTGC 620
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 Db 621 TACATGCCGAGACCCACGTTTCTCTCTCACTCAACACCAAGTACAGAGGCCATGGCT 680
 QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
 Db 681 GCCTTGGCTCTCTGTGGGCTCTGAGAGGGCTGGGAAGAGCCCTGTGTGGGCTGAG 740
 QY 241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIle 259
 Db 741 CACCAAGGCTTCCAGCTGGCCCTGTCTGAGGCGCTTGGCATCTACAAGCCCTCATCATC 800
 QY 260 GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla 279
 Db 801 GGCATTTCCCTCATGGTCTTCAGCAGCTGTCTGAGGGGTCAATGCTATCATGTTCTATGCC 860
 QY 280 GluThrIlePheGluAlaIlePheLysAspSerSerLeuLeuAspValValValGly 299
 Db 861 AACACATCTTCGAGGAGGCAAGTTCAAGGACAGCAGCCTGGCTGCTGCTGCTGCTGCTG 920
 QY 300 ValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArg 319

Db 921 ATAATCCAGGTCCTGTTCACTGCTGTGGCGCCCTCATCATGACAGACAGCGGCAAGG 980
 QY 320 LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr 339
 Db 981 CTGCTCTCCGCGCTTGTGGGTGTGATCATGTTGTTTCAGTATGAGTGCCTTTGGTACCTAC 1040
 QY 340 PheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProVal 359
 Db 1041 TTCAAACTGACCCAGAGCCTCCCCAGCAACTCCTCCACGCTAGGCTG---GTGCCCATC 1097
 QY 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValIleGlySerMetCys 379
 Db 1098 GCGCGGAGCCTGTGATGTCCAAGTGGAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1157
 QY 380 LeupheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGlu 399
 Db 1158 CTCTTCATGTGCTGCTTTCGCGTGGCTGGGACCACTCCCTGGCTCTCATGTGCAGAG 1217
 QY 400 IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuLeuThrAsnTrpLeu 419
 Db 1218 ATCTTCCCTCTGCATGTCAAGGCTGTGCTACCGCATCTGTGCTCTCACCACCTGGTTC 1277
 QY 420 MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuValLeuArgProTyrGly 439
 Db 1278 ATGCGCTTCTAGTGACCAAGAGTTCAACAGGCTCAAGAGTCAAGAGTCAAGAGTCAAGAGT 1337
 QY 440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
 Db 1338 GCCTTCTGGCTCACCGCTGCTCTGCGCTCTCAGTGTCTTATTCACACTGACCGTTGTC 1397
 QY 460 ProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 Db 1398 CCTGAGACTAAAGCAGGACTCTGGAACAAGTCAAGCCCATTTTCGAGGGACGA 1451
 RESULT 12
 AB033418
 LOCUS AB033418 2189 bp mRNA linear ROD 11-APR-2000
 DEFINITION Rattus norvegicus glut8 mRNA for glucose transporter 8, complete cds.
 ACCESSION AB033418
 VERSION AB033418.1 GI:7592743
 KEYWORDS Glucose transporter 8.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 2189)
 AUTHORS Ishibashi, K.
 TITLE Molecular cloning of a new putative glucose transporter
 JOURNAL Published Only in database (2000)
 REFERENCE 2 (bases 1 to 2189)
 AUTHORS Ishibashi, K.
 TITLE Direct Submission
 JOURNAL Submitted (08-OCT-1999) Kenichi Ishibashi, Jichi Medical School, Pharmacology, Minami-Kawachi, Kawachi, Tochigi 329-0498, Japan (E-mail:kishiba@jichi.ac.jp, Tel:81-285-58-7326, Fax:81-285-44-5541)
 FEATURES
 Location/Qualifiers
 1..2189
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone_lib="testis"
 1..2189
 /gene="glut8"
 109..1548
 /gene="glut8"
 /note="Glut8 has 12 transmembrane segments"
 /codon_start=1
 /product="glucose transporter 8"
 /protein_id="BAA94383.1"

/db_xref="GI:7592744"
/translation="WSPEDPOETOPLESFPGARAPGRRVFLATFAAALGPLSFGFAL
GYSPAIPLSLRTAPPALRLGDTAASWFGAVTILGAARQGVAGLVGSLDRAGRKLISLL
LCTPFTVGFVITPAARDVLMGLLGRLLTGACVSLVAPVYISEIAYPAVRGLIGS
CQVMVVTGIIILAYVAGVLEWRLAVLCVPPPTLMLLMCMYMPETPRFLITQHOYQE
AMALRFTWSEEGWEPVGAHQFOLAMLRPGVHKPLIIGICLMVFQOLSGVNA
IMFYANTPIFEAKFDSSLASVTVGIIQVLPFAVALIMDRAGRLKLLALSGVIMVFS
MSAFGTYPKLTQSGPNSSHVGLVPIISAEPADVHLGLAWLAVGSMCLPIAGFVGMG
DIPWLLMSEIIFPLHKIGVATGCVLTINFMFAPLVTKFNSIMEILRPGAFWLTAAPFC
ILSLVLTUTLTFVETKGRLEQITAHFEGR"

ORIGIN

Alignment Scores:
Pred. No.: 5, 7e-145 Length: 2189
Score: 2146.00 Matches: 407
Percent Similarity: 91.86% Conservative: 33
Best Local Similarity: 84.97% Mismatches: 37
Query Match: 87.34% Indels: 2
DB: 10 Gaps: 2

US-09-886-954a-1 (1-477) x AB033418 (1-2189)

QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
DB 109 ATGTGCGCGAGGACCCCGAGAAACACAGCGCGTGTTCGGGTGCGCGGCGCCAGGGCT 168
QY 21 ProArgGlyArgValPheLeuAlaPheAlaAlaLeuGlyProLeuSerPhe 40
DB 169 CCGCGCGCGCGCGGTCTCTTCCACCTTCGCGCGCGCGCTGGCGCGCGCTACGCTTC 228
QY 41 GlyPheAlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
DB 229 GGCCTTCGCGCTGGCTACAGCTCCCGCGCATCCCGAGCTCGGGCGCACCCCGCT 288
QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
DB 289 GCCCTACGCTCGGAGACACTCGGCGCTCTCGTTTCGGGCGCGCTGTGACCCCTGGCGCT 348
QY 81 AlaAla---GlyGlyValLeuGlyTyrLeuValAspArgAlaGlyArgGlyLeuSer 99
DB 349 GCGGCGCGGCGCGGTGTGCGGCGCTGGCTTCCTGGACCGTGCAGGGCGCAAGCTGAGC 408
QY 100 LeuLeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAsp 119
DB 409 CTCCTGCTCTGCACCGTGCCCTCTGTCGACCGGTTTGCTGTATCATCACCGCGCGCGGAT 468
QY 120 ValTrpMetLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeu 139
DB 469 GTGTGATGCTGCTCGAGCGCGCTCTCTCACCGCGCTAGCGCTGCGGAGTGCCTCATTA 528
QY 140 ValAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuGlySer 159
DB 529 GTGGGACCGGTCTATATCTCGGAATCGCTTACCAGCGCTTCGAGGACTGCTCGGCTCC 588
QY 160 CysValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeu 179
DB 589 TGTGTGACGCTGATGTTGTACTGTCATCTCTTGGCTCTACGTGCGAGGCTGGGTCCTTA 648
QY 180 GluTrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMet 199
DB 649 GAATGCGCTGTTGCTGTGTGGGCTGTGTGCGCGCGCTGTCGCGCGCTGCTGCTCATG 708
QY 200 CysPheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnAlaMet 219
DB 709 TGTACATGCCGAGACCCACGCTTCTCTTGACGCAACACAGTACCAGAGGCGCATG 768
QY 220 AlaAlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProLeuGlyAla 239
DB 769 GCTGCGCTCGCTTCTGTGGGCTCTGAGGAGGCGTGGGAAGAGCCCTCTGTGGGCT 828
QY 240 Glu---GlnSerPheHisLeuAlaLeuArgGlnProGlyIleTyrIleProPheIle 258
DB 829 GAGCACGAGGCTTCCAGCTGCTATGCTGAGGCGCGCTGGTGTCCACCAAGCCCTCATC 888

QY 259 IleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyr 278
DB 889 ATCGCATTTTCCTCATGGTCTTCCAGCAGCTGTCCAGGGTCAAGCCCATCATGTTCTAT 948
QY 279 AlaGluThrIlePheGluGluAlaLysPheLysPheSerSerLeuAlaSerValVal 298
DB 949 GCCAACACCATTTTGAGGAGGCCAAGTTCAAGGACAGCAGCGCTGGCTCGTCACTGTG 1008
QY 299 GlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArg 318
DB 1009 GGCATCATCCAGGTTCTCTTCACTGCTGTGGCGGCCCTCATCATGACAGACGAGCGGA 1068
QY 319 ArgLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAla 338
DB 1069 AAATGCTTCTGGCCTTGTGGGTGTGATCATGTTTCAGCATGAGCGCTTGGTACC 1128
QY 339 TyrPheLysLeuThrGlnGlyGlyProGlyAsnSerHisValAlaIleSerAlaPro 358
DB 1129 TACTTCAAACTGACCCAGAGTGGCCCGACCAACTCTCCCATGTAGGCTCTCGTGGTCC 1188
QY 359 ValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMet 378
DB 1189 ATCTCCGACAGCGCTGCTGATGTTTCACCTGGGCGCTGGCTGGCTGTAGGCGACATG 1248
QY 379 CysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSer 398
DB 1249 TGCCTTTCATCGCTGGTGTTCAGTAGCTGGGACCCATCCCTGGCTCTCATGTCA 1308
QY 399 GluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrp 418
DB 1309 GAGATCTTCCCTCTGCATCAAGGCTGTGGCTACCGGCGTGTGTCTCTACCAACATGG 1368
QY 419 LeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyr 438
DB 1369 TTCATGGCTTCTCTGTCACCAAGAGTTTAAACAGCATCATGAGATCTCTAGACCCCTAC 1428
QY 439 GlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCys 458
DB 1429 GCGCGCTTCTGGCTCACCGCTGCTGTATCTCTCAGCGCTCTTTCAGCGTCACTTTT 1488
QY 459 ValProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
DB 1489 GTCCCTGAGACTAAAGGAGGAGCTCTGGAACAAATCAGACCCATTTTCGAGGAGCGG 1545
RESULT 13
AX191507
LOCUS AX191507
DEFINITION Sequence 29 from Patent WO0149728.
ACCESSION AX191507
VERSION AX191507.1 GI:15209697
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Kato, S. and Kimura, T.
TITLE Human proteins having hydrophobic domains and dnas encoding these proteins
JOURNAL Patent: WO 0149728-A 29 12-JUL-2001;
PROTEGENE Inc. (JP) ; SAGAMI CHEMICAL RESEARCH CENTER (JP)
FEATURES
source
1. 1461
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
61. 849
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC51153.1"
/db_xref="GI:15209698"
/db_xref="RENTREMBL:CAC51153"

```
/translation="WTPEDPEETOPLLGPGGSGAPGRRVFLAAFAALGBLSPFGFAL
GYSPPSLQRAAPAPRPPDRDAASWFAVVTGAAGVGLGVDVDRAGRLKSLLL
CSVPFVAFPAITAAQDQVMMLGRLLTGLACVASVAPVTISETAPVAVGLGSS
VOLVMVGILLAYAGWLVEWRWLAVLGVCPVLLMLCMPEPTFELTQHRQREA
APGLVRGCGHVGQECRLLOADPGWVCPVLLRGHIGACITCTAC"
```

ORIGIN

Alignment Scores:

Pred. No.:	1.09e-102	Length:	1461
Score:	1554.50	Matches:	324
Percent Similarity:	67.85%	Conservative:	1
Best Local Similarity:	67.64%	Mismatches:	1
Query Match:	63.21%	Indels:	153
DB:	6	Gaps:	2

US-09-886-954A-1 (1-477) x AX191507 (1-1461)

QY	1	MetThrProGluAspProGluGlnThrGlnProLeuLeuGlyProProGlyGlySerAla	20
Db	61	ATGACCCCGAGACCCAGAGAAACCCAGACCGCTTCTGGGCGCTCTCTGGCGGCGAGCGC	120
QY	21	ProArgGlyArgArgValPheLeuAlaIaPheAlaAlaLeuGlyProLeuSerPhe	40
Db	121	CCCCGGCGCGCGCGGTCTTCTCGCGGCTTCGCCGTGCGCTGGGCCCACTCAGCTTC	180
QY	41	GlyPheAlaLeuGlyTyrSerSerProLalleProSerLeuGlnArgAlaAlaProPro	60
Db	181	GGCTTCGCGCTCGGCTACAGCTCCCGGCATCCCTAGCTGCAGCGCGCGGCCCGC	240
QY	61	AlaProArgLeuAspAlaAlaIaSerTrpPheGlyAlaValValThrLeuGlyAla	80
Db	241	GCCCGCGCTGGAGCAGCGCCGCCCTCTCGTTCGGGCTGCTGACCCCTGGGTGCC	300
QY	81	AlaAlaGlyGlyValLeuGlyGlyTyrLeuValAspArgAlaGlyArgLysLeuSerLeu	100
Db	301	CGCGGGGGGAGTGTGGCGGCTGGCTGGTGGACCGCGCGGCGCCAGCTGAGCCTC	360
QY	101	LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal	120
Db	361	TTCGTGTGCTCGCTGCCCTTCGTGCCCGGCTTGCCTCATCACCGCGGCCAGGACGTG	420
QY	121	TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaAlaSerLeuVal	140
Db	421	TGGATGCTCTGGGGGGCGGCTCTCTACCGGCTGGCTGCGGTGTGGCTCCCTAGTG	480
QY	141	AlaProValTyrIleSerGluLeuAlaTyrProAlaValArgGlyLeuLeuGlySerCys	160
Db	481	GCCCGGCTACATCTCCGAAATCGCTTACCAGCAGTCCGGGGGTGTGCTCGCTCCTGT	540
QY	161	ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGlu	180
Db	541	GTGCAGCTAATGGTCGTGCGCATCTCTCGCTACCTGGCAGCTGGGTGCTGGAG	600
QY	181	TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys	200
Db	601	TGGCGCTGCGCTGCTGCTGGCTGCTGCGCCCTCCCTCATGCTGCTTCTCATGTGC	660
QY	201	PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAla-MetAl	220
Db	661	TTCATGCCCGAGACCCCGCTTCTGTGACTCAGCACAGGCGCCAGGAGGTGCTC--	718
QY	220	aAlaLeuArgPheLeuTyrGlySerGluGlnGlyTyrGluAspProIleGlyAlaG	240
Db	718	-----	718
QY	240	uGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIysProPheIleIleG	260
Db	718	-----	718
QY	260	yValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaG	280
Db	718	-----	718

Qy	280	uThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValcIyVa	300
Db	718	-----	718
Qy	300	lIleGlnValLeuPheThrAlaValAlaAlaIleuIleMetAspArgAlaGlyArgArgIe	320
Db	718	-----	718
Qy	320	uLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyValatyrPh	340
Db	719	---CTGGTCITGTTCAGGTGTGGTCATGGTTCAGCACGAGTGCCTTCGGCGCCTACTT	774
Qy	340	eLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSe	360
Db	775	CAAGCTGACCCAGGGTGCCCTGGCAACTCCTCGCACGTGGCCATCTCGGGCCCTGTCTC	834
Qy	360	rAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysIe	380
Db	835	TGCACGCTGTTCATGCCACGCTGGGGCTGGCTGGCTGGCGTGGGACGATGTGCTT	894
Qy	380	uPheIleAlaGly-PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluI	400
Db	895	CTTCATCGCCGGAGTC-----	911
Qy	400	lePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuM	420
Db	911	-----	911
Qy	420	etAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyVa	440
Db	912	-----CTCAGGCCCTATGGAG	927
Qy	440	laPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValP	460
Db	928	CCITCTGGCTTGCCTCCGCTTCTGCATCTTCAGTGTCTCTTTCACCTTTGTCTGTGCC	987
Qy	460	roGluThrLysGlyLysThrLeuGlnIleThrAlaHisPheGluIyArg	477
Db	988	CTGAACCTAAAGGAAAGAACTCTGGAACAAATTCACAGGCCATTTTTCAGGGGGA	1040

RESULT 14
AF321324

AF221324	LOCUS	AF221324	1012 bp	mRNA	linear	MAM 02-JUL-2001
DEFINITION	Eos taurus glucose transporter 8 (Glut8)			mRNA	partial cds.	

ACCESSION	AF321324
VERSION	AF321324.1
	GI:14582715

KEYWORDS
SOURCE
Bos taurus (cow)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovinae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 1012)

AUTHORS Augustin, R., Navarrete-Santos, A. and Fischer, B.

TITLE	Direct Submission

JOURNAL Submitted (13-NOV-2000) Anatomy and Cell Biology, Faculty of

Medicine, Martin Luther University Halle, Grosse Steinstrasse 52,

Halle 061097, Germany

FEATURES
source

/organism="Bos taurus"

/mol type="mrna"

/db_xref=

gene <1. .1012

```
/gene="Glut8"
```

CDS <1. .1005

```

/gene="Glut8"

```

/codon_start=

```

/product="glucose transporter 8"

```

```
/protein_id="AAK69606.1"
```

/db_xref="GI:14582716"

```
/translation="ISEIAYPEVRGLLGSCVQLMVVTGILLAYLAGWVLEWRWLAVLG
```

AOLRRPGVYKPIIGISLMAFQOLSGNVAMVFAETIPEEKFKDSSLASVVVGVIQV
LFTTAALIMDRAGBRLLLTISGVVMVFTSARFTYFKLTGGSPNSHVDLPALVSM
EADNTNVLAMLVGNMCLPIAGFAVWGPIFWLLMSLFPPLHVKGVATGVCYVLNWF
MAFLVTKFSSLMELVRPYGAFWLASAFICFVLEFLACVPEIKTKLTLEQITAHPEGR
"

ORIGIN

Alignment Scores:
Pred. No.: 1,17e-101 Length: 1012
Score: 1537.50 Matches: 293
Percent Similarity: 92.81% Conservative: 17
Best Local Similarity: 87.72% Mismatches: 23
Query Match: 62.58% Indels: 1
DB: 4 Gaps: 1

US-09-886-954A-1 (1-477) x AF321324 (1-1012)

Qy 145 llsSerGluileAlaTyProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMet 164
Db 1 ATCTCGAAATCGCTACCTGAGGTGGAGGGCTCGGGCTCTGTGTGCAGCTGATG 60
Qy 165 ValValValGlyLeuLeuAlaTyTrpValGlnGlyTrpValLeuGluTrpArgTrpLeu 184
Db 61 GTAGTCACAGGCATCTCTACCTACCTGGCAGGCTGGGTCTCGAGTGGCGCTGGCTG 120
Qy 185 AlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMetProGlu 204
Db 121 GCCGTGCTGGGTGTGGCCCTCTTCATGCTGCTGCTCATGTGTTTCATGCTGAG 180
Qy 205 ThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPhe 224
Db 181 ACCCTCGCTTCTCTGCTGCTCAGCACACACAGGAGGCCATGGCCGCCATCGATTTC 240
Qy 225 LeuTrpGlySerGluGlnGlyTrpGluaspProproIleGlyAlaGlu---GlnSerPhe 243
Db 241 CTGTGGGGCTATGCCAGGGTTGGAGGAGCCCTCCCTTGGGGCTCAGCACAGGACTTC 300
Qy 244 HisLeuAlaLeuLeuArgGlnProGlyLeuTyLysProPheIleIleGlyValSerLeu 263
Db 301 CAGGTGGCCAGCTGGGGGTCTGTGTGTCACAGCCCTTCATATCGGCATCTCCCTG 360
Qy 264 MetAlaPheGlnLeuSerGlyValAsnAlaValMetPheTyAlaGluThrIlePhe 283
Db 361 ATGGCTTCCAGCAGCTGTGAGCGGTCAATGCGCTCATGTTCTACGCGAGACCATCTTT 420
Qy 284 GluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyValIleGlnVal 303
Db 421 GAGGAGGCTTAAGTTCAAGACAGCAGCCCTGGCTCAGTCGTGTGGGGCTCATCCAGGTG 480
Qy 304 LeuPheThrAlaValAlaAlaLeuLeuMetAspArgAlaGlyArgArgLeuLeuVal 323
Db 481 CTGTTCACTGCCACGGGGCTTCATGATGGACAGCTGGCGAAGGTGCTCTTGACC 540
Qy 324 LeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyTrpPheLysLeuThr 343
Db 541 TTGTCAAGTGTGTGTGATGTTTCAGCACCGCGCTTTGGTACCTACTTCAAGCTGACC 600
Qy 344 GlnGlyGlyProGlyAsnSerHisValAlaIleSerAlaProValSerAlaGlnPro 363
Db 601 GAGGGGGGGCCAGCAACTCTCCACGTGGACCTCCCGGGCTGCTCTCCATGGAGGCC 660
Qy 364 ValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAla 383
Db 661 GCCGATACCACAGTGGGGCTGGCTGGCTGGCGGTGGGCAACATGTGCTCTTCATCGCC 720
Qy 384 GlyPheAlaValGlyTrpGlyProIleProTrpTrpLeuLeuMetSerGluIlePheProLeu 403
Db 721 GGCTTCGCTGTGGGGTGGGGGCCATCCCGTGGCTCTCTCATGTGTCAGAGATCTTCCCTG 780
Qy 404 HisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeu 423
Db 781 CATGTCAAGGGCGTGGCCACCGCGCTGCTGGCTCTCACCACACTGGTTCATGGCTTCTG 840

Qy 424 ValThrLysGluPheSerSerLeuMetGluValLeuArgProTyTrpGlyAlaPheTrpLeu 443
Db 841 GTGACCAAGAGTTCAGAGTCTCATGAGGTGCTCAGGCCCTACGGTTCCTTCTGGCTG 900
Qy 444 AlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrLys 463
Db 901 GCCTCTGCCTTCTGCATCTCGGTGCTCTTTTCACTTGGCTGTGTCTCCCTGAGACCAA 960
Qy 464 GlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 961 GGGAGACTCTGGAAACAATCACAGCCCATTTTGGGGCCGA 1002
RESULT 15
BC049409
LOCUS 1929 bp mRNA linear VRT 07-OCT-2003
DEFINITION Danio rerio cDNA clone MGC:56364 IMAGE:5604396, complete cds.
ACCESSION BC049409
VERSION BC049409.1 GI:29436433
KEYWORDS MGC.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 1929)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaefer, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Hellon, E., Kettman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalusz, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
REFERENCE 2 (bases 1 to 1929)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Sumio Sugano
cDNA Library Preparation: Dr. Sumio Sugano
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzyzanski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalusz, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

QY	2	ThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAlaPro	21
Db	197	ACAGAGAGGATAGATCCGAGCAAGACGCATACCTG-----GATAAAGTT	241
QY	22	ArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGly	41
Db	242	AGAAATGGAAAGCTGTTATCGCCACCTTCGCAGCAGTTCTGGAGACGCTTAGTTTGGG	301
QY	42	PheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProAla	61
Db	302	TTTCGTGCTGGCTACAGCTCTCCTGCTATCCCTCAACTCCGCAGATCCCAAGACCTGAGA	361
QY	62	ProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAla	81
Db	362	CTCAGCTCAGCGTACAGAGGAAGTTCCTGTTGCGGTCTGCTGTGACCATTCGTGGCGCT	421
QY	82	AlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgIstyeuSerLeuLeu	101
Db	422	CTTGGGGGTCCTCTGGAGAGATGATTTGGAGAGGATCGGACAAAGTTTGACCCCTATG	481
QY	102	LeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrp	121
Db	482	TTCTGTGCATACCATTCATCTTTGGGTTTACCACCATTATCGCTGCTCAGACCACTGG	541
QY	122	MetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAla	141
Db	542	ATGTTCTATGTGGCAGAGTGCCTCATGGTTTGGCAGTCGATGCCTCTTAGTTGTA	601

Search completed: September 28, 2004, 03:28:36
Job time : 4617 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 28, 2004, 03:28:39 ; Search time 4236 Seconds
(without alignments)
570.345 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPEDPBTQPLLPQGS.....CVPETKGLTLEQTAHEGR 477

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 337386 seqs, 2532474692 residues

Total number of hits satisfying chosen parameters: 6674772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-Q=/cgn2_1/USFPO_spool_p/US09886954/runat_27092004_170634_13496/app_query.fasta_1.647
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DLOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09886954@cgn_1_1_781_runaat_27092004_170634_13496
-NCFU=6 -ICFU=3 -NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :			
Published Applications NA.*			
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*			
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*			
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*			
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*			
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*			
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*			
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*			
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*			
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*			
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*			
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*			
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*			
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*			
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*			
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*			
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*			
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*			
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*			
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*			
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.			
SUMMARIES			
Result No.	Score	Query Match Length	Description

ALIGNMENTS			
RESULT 1			
US-10-115-831-137			Sequence 137, App
Sequence 137, Appl			Sequence 28, Appl
Sequence 138, App			Sequence 138, Appl
Sequence 688, App			Sequence 688, Appl
Sequence 29, Appl			Sequence 29, Appl
Sequence 19, Appl			Sequence 19, Appl
Sequence 358, App			Sequence 358, Appl
Sequence 1685, App			Sequence 1685, Appl
Sequence 109, App			Sequence 109, Appl
Sequence 13608, A			Sequence 13608, A
Sequence 132422, A			Sequence 132422, A
Sequence 26596, A			Sequence 26596, A
Sequence 61061, A			Sequence 61061, A
Sequence 24968, A			Sequence 24968, A
Sequence 26550, A			Sequence 26550, A
Sequence 35, Appl			Sequence 35, Appl
Sequence 357, App			Sequence 357, App
Sequence 820, App			Sequence 820, App
Sequence 64291, A			Sequence 64291, A
Sequence 47058, A			Sequence 47058, A
Sequence 22342, A			Sequence 22342, A
Sequence 23817, A			Sequence 23817, A
Sequence 74980, A			Sequence 74980, A
Sequence 47223, A			Sequence 47223, A
Sequence 22366, A			Sequence 22366, A
Sequence 22036, A			Sequence 22036, A
Sequence 6734, App			Sequence 6734, App
Sequence 47239, A			Sequence 47239, A
Sequence 1627, App			Sequence 1627, App
Sequence 2331, App			Sequence 2331, App
Sequence 38732, A			Sequence 38732, A
Sequence 40124, A			Sequence 40124, A
Sequence 1769, App			Sequence 1769, App
Sequence 96711, A			Sequence 96711, A
Sequence 79, Appl			Sequence 79, Appl
Sequence 35021, A			Sequence 35021, A
Sequence 507, App			Sequence 507, App
Sequence 1341, App			Sequence 1341, App
Sequence 1341, App			Sequence 1341, App
Sequence 248, App			Sequence 248, App
Publication No. US20030219743A1 Nucleic Acids and			
GENERAL INFORMATION:			
APPLICANT: Tang, Y. Tom			
APPLICANT: Liu, Chenchua			
APPLICANT: Asundi, Vinod			
APPLICANT: Ren, Feiyan			
APPLICANT: Drmanac, Radoje T.			
TITLE OF INVENTION: No. US20030219743A1			
FILE OF INVENTION: Polypeptides			
FILE REFERENCE: 792CIP2ADIV			
CURRENT APPLICATION NUMBER: US/10/115,831			
PRIORITY FILING DATE: 2002-04-02			
PRIORITY FILING DATE: 09/667,298			
PRIORITY FILING DATE: 2000-09-22			
PRIORITY FILING DATE: 09/577,408			
NUMBER OF SEQ ID NOS: 178			
SOFTWARE: pt_FL_genes Version 2.0			
SEQ ID NO 137			
LENGTH: 1862			

;
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: {4}..(1437)
US-10-115-831-137

Alignment Scores:

Pred. No.: 3e-254 Length: 1862
Score: 2457.00 Matches: 477
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

US-09-886-954A-1 (1-477) x US-10-115-831-137 (1-1862)

Qy 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlySerAla 20
Db 4 ATGACGCCGAGGAGCCAGAGAAACCCAGCGCTTCTGGGGCTTCTGGCGCAGCGC 63
Qy 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
Db 64 CCCCGCGCGCGCGCTTCTCGCCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTT 123
Qy 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaPro 60
Db 124 GCTTCGCGCTCGGCTCAGCTCCCGGCATCCCTAGCTGACGCGCGCGCGCGCG 183
Qy 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
Db 184 GCCCGCGCGCTGGACGCGCGCGCTTCTGGTTTCGGGGCTGCTGACCCGCGTGC 243
Qy 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlySerLeu 100
Db 244 GCGCGCGCGGAGTCTCGCGCGCTTCTGGCGCGCTTGGCGCTTGGCGCTTGGCG 303
Qy 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db 304 TTGCTGTGCTCGTGCCTTCTGGCGCGCTTGGCGCTTGGCGCTTGGCGCTTGG 363
Qy 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db 364 TGGATGCTGCTGGGGCGCGCTCCTCAGCGCGCTGCGCTGCGGTGTTGCTCCCTAGTG 423
Qy 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
Db 424 GCCCGCGCTACATCTCGAAATCGCTACCCAGCAGTCCCGGGGTGCTCGGCTCCTGT 483
Qy 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
Db 484 GTGCAGCTAATGCTGCTCGGCATCTCTGGCTACTTGGCAGGCTGGGTGCTGGAG 543
Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 544 TGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 603
Qy 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
Db 604 TTCATGCCCGAGACCCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
Qy 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProPheGlyAlaGlu 240
Db 664 GCCCTGGGTTCCTGTGGGCTCCGAGCAGGCTGGAGAGACCCCGCATCGGGCTGAG 723
Qy 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIleProPheIleleGly 260
Db 724 CAGAGCTTTACCTGGCGCTGCTGGCGCAGCGCGCGCATCTACAAGCCCTTCATCATCG 783
Qy 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
Db 784 GTCTCCCTGATGGCTTCCAGCAGCTGCTGGGGGTCAAGCGCTATGTTTCATGACAG 843

Qy 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyVal 300
Db 844 ACCATCTTGAAGAGGCCAAGTTCAGGACAGCAGCTGGCTCGGTCTGGTGGGTGC 903
Qy 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeu 320
Db 904 ATCCAGGTGCTGTTTCACAGCTGTGCGGCTCTCATCATGAGACAGAGCGGAGGCTG 963
Qy 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
Db 964 CTCCTGGTCTGTGCTGAGGTGTGCTCATGTGTTCAGCAGAGTGGCTTCGGCGCTTCT 1023
Qy 341 LysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
Db 1024 AAGCTGACCCAGGGTGGCCCTGGCAACTCTCGCAGCGTGGCACTCTCGGCGCTGTCT 1083
Qy 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
Db 1084 GCACAGCCTGTGTATGCCAGCGTGGCGCTGGCTGGCTGGCGGAGCATGTGCGCTC 1143
Qy 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400
Db 1144 TTCATCGCGCGCTTTCGCTGGGTGGCGGCCATCCCTGCTGCTCTCATGTTCAGAGATC 1203
Qy 401 PheProLeuHisValLysGlyValAlaAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
Db 1204 TTCCTCTGCATGTCAAGGGCGTGGCGACAGGCACTCTCGCTCTCAACCACTGGCTCATG 1263
Qy 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
Db 1264 GCCTTCTCTGTGACCAAGAGTTCAGCAGCCTCATGGAGGTCTCTCAGGCGCTATGGAGCC 1323
Qy 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
Db 1324 TTCGGCTTGGCTCGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 1383
Qy 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 1384 GAAACTAAAGGAAAGACTCTGGAAACAATCACAGCCCATTTTGGGGCGGA 1434

RESULT 2

US-10-168-651-28
; Sequence 28, Application US/10168651
; Publication No. US20030171275A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BURFORD, Neil
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: YANG, Junming
; APPLICANT: REDDY, Roopa
; APPLICANT: LAL, Preeti
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Valda
; APPLICANT: YUE, Henry
; APPLICANT: NGUYEN, Dannie B.
; APPLICANT: YAO, Monique G.
; APPLICANT: GANDHI, Ameen R.
; APPLICANT: TANG, Y. Tom
; APPLICANT: KHAN, Farrah A.
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
; FILE REFERENCE: PI-0005 PCT
; CURRENT APPLICATION NUMBER: US/10/168,651
; PRIOR APPLICATION NUMBER: 2002-06-21
; PRIOR FILING DATE: 60/172,000; 60/176,083; 60/177,332; 60/178,572; 60/179,758
; PRIOR FILING DATE: 1999-12-23; 2000-01-14; 2000-01-21; 2000-01-28; 2000-02-02;
; 2000-02-10
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 28
; LENGTH: 2080

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030171275A1 1416107CB1
US-10-168-651-28

```

Alignment Scores:

```

Pred. No.: 3,51e-254 Length: 2080
Score: 2457.00 Matches: 477
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

```

US-09-886-954A-1 (1-477) x US-10-168-651-28 (1-2080)

```

Qy 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db 46 ATGACGCCGAGGAGACCAGAGAAACCCAGCCGCTTCTGGGGCTCTCTGGCGGAGCGCG 105
Qy 21 ProArgGlyArgValPheLeuAlaPheAlaAlaLeuGlyProLeuSerPhe 40
Db 106 CCCCAGCGCGCGCGCGCTTCTCTCGCGCGCTTCTCGCGCGCTTCTCGCGCGCGCGCGCTTCT 165
Qy 41 GlyPheAlaLeuGlyTyrSerSerProAlaTleProSerLeuGlnArgAlaAAlaProPro 60
Db 166 GGCTTCGGGCTCGGCTACAGCTCCCGCGCCATCCCTAGCTGACGCGCGCGCGCGCGCGCG 225
Qy 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValAlaValThrLeuGlyAla 80
Db 226 GCCCGCGCGCTGGACGAGCGCGCGCTCTCTGGTTCGGGGCTGTCGTGACCTGGGTGCC 285
Qy 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db 286 GCGCGCGCGGAGTGCCTGGGCGGTGGCTGGTGACCGCGCGCGCGCGCGCGCGCGCGCGCTC 345
Qy 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValleThrAlaAlaGlnAspVal 120
Db 346 TTGCTGTGCTCCGTGCGCTTCGTGGCGGCTTCGTGCGGCTTCATCCCGCGCGCGCGCGCG 405
Qy 121 TrpMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaAlaSerLeuVal 140
Db 406 TGGATGTGTGGGGGCGCGCTCTCTCACGGCTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 465
Qy 141 AlaProValTyrIleSerGluAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
Db 466 GCGCGCGCTTACATCTCCGAAATCGCCTACCCAGCAGTCCGGGGGTTGCTCGGCTCCTGT 525
Qy 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
Db 526 GTGAGCTATGTGCTGCGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCT 585
Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 586 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 645
Qy 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
Db 646 TCCATGCCCGAGACCCCGCGCTTCTCTGCTACTCAGCACAGCGCGCGCGCGCGCGCGCGCG 705
Qy 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
Db 706 GCCCTGGGTTCTCTGGGCTCCAGCAGGCTGGGAAGACCCCGCGCGCGCGCGCGCGCGCG 765
Qy 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGly 260
Db 766 CAGAGCTTTCACCTGGCGCTGCTGGCGAGCGCGCGCGCTACAGCGCGCTTTCATCATCGCG 825
Qy 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
Db 826 GTCTCCCTGATGGCTTCCAGCAGCTGCGGGGTCAACGCGGTCACTGTTCTATGACAGAG 885

```

```

Qy 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyVal 300
Db 886 ACCATCTTTGAAGAGGCCAAGATTCAAGACAGCAGCCTGGCTCGTCTGCTGGTGC 945
Qy 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeu 320
Db 946 ATCCAGGTGCTTTCACAGCTGGCGGCTCTCATCATGACACAGAGGAGGCGGAGGCTG 1005
Qy 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
Db 1006 CTCCTGGTCTTGTTCAGGTGGTTCATGGTTCAGCACGAGTGGCTTCGGCGCTACTTC 1065
Qy 341 LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
Db 1066 AGCTGACCCAGGCTGGCCCTGGCACTCTCTCGCACGTGGCCATCTCGGCGCTGTCTCT 1125
Qy 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
Db 1126 GCACAGCTGTTGATGCCAGCGTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTG 1185
Qy 381 PheIleAlaGlyPheAlaValGlyTyrGlyProIleProTrpLeuLeuMetSerGluIle 400
Db 1186 TTCATCGCGCGCTTGGGTGGGCTGGGCGGCCCATCCCTGGCTCTCATGTACAGATC 1245
Qy 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
Db 1246 TTCCTCTGCTGATGCAAGGCGTGGCGACAGGATCTGCGTCTCACCACTGGCTCATG 1305
Qy 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
Db 1306 GCCTTCTGCTGACCAAGAGTTCAGCAGCTCATGGAGGTCTCAGGCGCTATGGAGCC 1365
Qy 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
Db 1366 TTCCTGGCTTGGCTGGCTTCTGCTATCTCAGTCTCTTCTTCTTCTTCTTCTTCTTCT 1425
Qy 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 1426 GAACTAAAGAAAGACTCTGGAACAAATCACGCCCAATTTTGAGGGGCGA 1476

```

RESULT 3

```

US-10-115-831-138
; Sequence 138, Application US/10115831
; Publication No. US20030219743A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219743A1e1 Nucleic Acids and
; FILE REFERENCE: 792CIP2ADIV
; CURRENT APPLICATION NUMBER: US/10/115,831
; PRIOR FILING DATE: 2002-04-02
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/667,298
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 138
; LENGTH: 1655
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4)..(1230)
US-10-115-831-138

```

```

Alignment Scores: 2,25e-212 Length: 1655
Pred. No.: 2067.50 Matches: 408
Score:

```

Percent Similarity: 85.53% Conservative: 0
Best Local Similarity: 85.53% Mismatches: 0
Query Match: 84.15% Indels: 69
DB: 16 Gaps: 1

US-09-886-954A-1 (1-477) x US-10-115-831-138 (1-1655)

QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
DB 4 ATGACGCCCGAGGACCCAGAGAAACCCAGCCGCTTCGCGGCGCTTCGCGGCGAGCGCG 63
QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
DB 64 CCCCGCGCGCGCGCTTCCTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCTTC 123
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
DB 124 GGCTTCGCGCTCGGCTACAGCTCCCGCGCCATCCCTAGCCTGCAGCGCGCGCGCGCGCG 183
QY 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
DB 184 GCCCGCGCGCTGAGACGCCCGCGCTTCCTCGGTTGCGG----- 222
QY 81 AlaAlaGlyGlyValLeuLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
DB 222 ----- 222
QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
DB 222 ----- 222
QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
DB 222 ----- 222
QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
DB 223 -----GTCATACATCCGAATCGCTACCCAGCAGTCCCGGGGTTCGCGCTCGT 276
QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
DB 277 GTGAGCTAATGGTCGTGCTCGGCATCCCTCGCGCTACCTGGCAGGCTGGGTGTGGAG 336
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
DB 337 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 396
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
DB 397 TTCATGCCGAGACCCCGCGCTTCCTGCTGACTCAGCACAGCGCGCGCGCATGGCC 456
QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
DB 457 GCCCTCGGCTTCCTGCGGCTCCGAGCAGGCGTGGGAAGACCCCGCATCGCGGCGTGA 516
QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGly 260
DB 517 CAGAGCTTTCACCTGGCCCTGCTGGCGAGCCCGGCATCTACAGCCCTTCATCATCGGT 576
QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
DB 577 GTCTCCCTGATGGCTTCAGCAGCTGTCGCGGGTCAACGCGCTCATCTTCATGCAGAG 636
QY 281 ThrIlePheGluAlaAlaLysPheLysAspSerSerLeuAlaSerValValValGlyVal 300
DB 637 ACCATCTTTGAAGAGGCAAGTCAAGACAGAGCCCTGGCGCTCGGTCTCGTGGGTGTC 696
QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeu 320
DB 697 ATCCAGGTGCTGTTACAGCTGTGGCGCTCTCATCATGGACAGCAGCGCGGAGGCTG 756
QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrAlaPheGlyAlaTyrPhe 340
DB -----

757 CTCCTGGTCTTGTTCAGGTGTGGTTCATGGTGTTCAGCAGAGTTCCTTCGGCGCCTACTTTC 816
QY 341 LysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
DB 817 AAGCTGACCCAGGTTGGCCCTCGCAACTCTCGACGTCGGCCATCTCGGCGCCTGTCTCT 876
QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
DB 877 GCACAGCGCTGTGTATGTCAGCGTGGGCTGGCTGGCTGGCTGGCGGCGAGCATGTGCGTC 936
QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProThrLeuLeuMetSerGluIle 400
DB 937 TTCATCGCGCTTTCGGGTGGGCTGGGCGCCATCCCTGGCTCCTCATGTCAAGATC 996
QY 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
DB 997 TTCCTCTGATGTCAAGGGCGTGGCAGAGCATCTCGCTCCATCCCAACTGGCTCATG 1056
QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
DB 1057 GCCTTTCCTGTCACCAAGGAGTTCAGCAGCTCATGGAGTCTCAGGCCCTATGGAGCC 1116
QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
DB 1117 TTCGGCTTGGCTCCGCTTTCGTCATCTTCAGTGTCTCTTTCATCTTGTCTGTGCTCT 1176
QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
DB 1177 GAACTAAAGAAAGACTCTGGAACAAATCACAGCCCATTTTGGGGGCGGA 1227

RESULT 4
US-10-264-237-688
; Sequence 688, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 688
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (24)..(24)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (351)..(351)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (397)..(397)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (432)..(432)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (792)..(792)

OTHER INFORMATION: n equals a,t,g, or c
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (1119)..(1119)
 OTHER INFORMATION: n equals a,t,g, or c
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (1156)..(1156)
 OTHER INFORMATION: n equals a,t,g, or c
 US-10-264-237-688

Alignment Scores:
 Pred. No.: 8,28e-171 Length: 1156
 Score: 1679.50 Matches: 344
 Percent Similarity: 72.12% Conservative: 0
 Best Local Similarity: 72.12% Mismatches: 10
 Query Match: 68.38% Indels: 125
 DB: 16 Gaps: 1

US-09-886-954A-1 (1-477) x US-10-264-237-688 (1-1156)

QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyCysSerAla 20
 DB 53 ATGAGCCCGAGGACCCAGAGAAACCCAGCGCTTCTGGGGCCCTCTGGGGCCAR---- 108
 QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
 DB 108 ----- 108
 QY 41 GlyPheAlaLeuGlyTy-SerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 DB 108 ----- 108
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 DB 108 ----- 108
 QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 DB 108 ----- 108
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 DB 108 ----- 108
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 DB 108 ----- 108
 QY 141 AlaProValTrpIleSerGluIleAlaValProAlaValArgGlyLeuLeuGlySerCys 160
 DB 109 -----GTCATACATCCGAAATCCCTACCCAGCAGTCGGGGGTGCTCGGCTCTGT 162
 QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaValTrpValLeuGlu 180
 DB 163 GTGCAGTAATGTCGTCGTCGGCATCTCTCGGCTACCTGGCAGGCTGGTCTGGAG 222
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 DB 223 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 282
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 DB 283 TTCAATGCCGAGACCCCGCGCTTCTGCTGACTCAGCAGCGCCAGGAGGCGCATGGCC 342
 QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
 DB 343 GCCCTGGGNTTCTGTGGGCTCCGAGCAGGCTGGGAAGACCCCCCATCGGNCCTGAG 402
 QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTrpLysProPheIleLeuGly 260
 DB 403 CAGAGCTTTACCTGGCGCTGCTGGCGCAGCCCGGCATCTACAAGCCCTTCAATCATCGGC 462

QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTrpAlaGlu 280
 DB 463 GTCTCCCTGATGGCCTTCCAGCAGCTGTCGGGGGTCAACGCCGTCTATGTTCTATGCAGAG 522
 QY 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyVal 300
 DB 523 ACCATCTTTGAGAGGCCCAAGTTCAAGACACAGCCTGGCCTCGTCTCGTGGGTGTC 582
 QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeu 320
 DB 583 ATCCAGGTGCTGTTTCCAGCTGTGGCGCTCTCATCATGGACAGAGCGGCGGCTG 642
 QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTrpPhe 340
 DB 643 CTCTCTGGTCTTGTTCAGGTGTGTCATGTTTCAGCAGAGTGCCTTCGGCGCCTACTTC 702
 QY 341 LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
 DB 703 AAGCTGACCCAGGCTGGCCCTGGCACTCTCGCACGTGGCCATCTCGGCGCTGTCTCT 762
 QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
 DB 763 GCACAGCCTGTGATGCCAGCGTGGGGCTNGCCTGGCTGGCGTGGGCGAGCATGTGCCTC 822
 QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400
 DB 823 TTCAATCCCGCGCTTTCGGGTGGGCTGGGGGCCCATCCCTCGCTCATGTCCAGAGATC 882
 QY 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
 DB 883 TTCCCTTGTGATGTCAGAGGCGTGGCAGCAGCAWYTCGCTCCTCACAACACTGGCTCATG 942
 QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTrpGlyAla 440
 DB 943 GSCTTTCTYKGACCAAGAGTTTCAGCAGCCTCATGGAGGTCTCTCAGGCCCTATGGAGCC 1002
 QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 DB 1003 TTTCTGGCTTGCCTCCGCTTTCGTCATCTTCAGTGTCTCTTCACTTTGGTCTGTGC-CCT 1061
 QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 DB 1062 GAAACTAAGAAAGAACTCTGGAACAATCAGACCCATTTTGGAGGGCGA 1112

RESULT 5

US-10-169-395-29
 ; Sequence 29, Application US/10169395
 ; Publication No. US20040034192A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KATO, Seishi
 ; APPLICANT: KIMURA, Tomoko
 ; TITLE OF INVENTION: HUMAN PROTEINS HAVING HYDROPHOBIC DOMAINS AND DNAS ENCODING
 ; TITLE OF INVENTION: THESE PROTEINS
 ; FILE REFERENCE: 01997, 015100, US
 ; CURRENT APPLICATION NUMBER: US/10/169,395
 ; CURRENT FILING DATE: 2002-11-29
 ; PRIOR APPLICATION NUMBER: JP 2000-585
 ; PRIOR FILING DATE: 2000-01-06
 ; PRIOR APPLICATION NUMBER: JP 2000-588
 ; PRIOR FILING DATE: 2000-01-06
 ; PRIOR APPLICATION NUMBER: JP 2000-2299
 ; PRIOR FILING DATE: 2000-01-11
 ; PRIOR APPLICATION NUMBER: JP 2000-26862
 ; PRIOR FILING DATE: 2000-02-03
 ; PRIOR APPLICATION NUMBER: JP 2000-58367
 ; PRIOR FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: PCT/JP00/09359
 ; PRIOR FILING DATE: 2000-12-28
 ; NUMBER OF SEQ ID NOS: 150
 ; SEQ ID NO 29
 ; LENGTH: 1461
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(849)
US-10-169-395-29

Alignment Scores:
Pred. No.: 3,34e-157 Length: 1461
Score: 1554.50 Matches: 324
Percent Similarity: 67.85% Conservative: 1
Best Local Similarity: 67.64% Mismatches: 1
Query Match: 63.27% Indels: 153
DB: 13 Gaps: 2

US-09-886-954A-1 (1-477) x US-10-169-395-29 (1-1461)

Qy 1 MetThrProGluAspProGluGlnThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db 61 ATGACGCCCGAGGACCCAGAGAAACCCAGCGCTTCCTGGGGCCCTCTGGCGCAGCGCG 120
Qy 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
Db 121 CCCCGCGCGCCCGCTTCCTCGCCGCTTCGCGGCTCGCCCTGGGCGCCACTCAGCTTC 180
Qy 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
Db 181 GGCCTCGCGCTCGGCTCAGCTCCCGGCGATCCCTAGCTCGACGGCGCCGCCGCCCG 240
Qy 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
Db 241 GCCCGCGCGCTGGACGCGCGCCCTCTGGTTCGGGGCTGTGCTGACCTCGGTGCTGC 300
Qy 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlyLeuSerLeu 100
Db 301 GCGCGCGGGGAGTGCTGGCGGCTGTGGTGTGGACCGCGCGCGCGCAAGCTGAGCTTC 360
Qy 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db 361 TTGCTGTGCTCCGTCGCCCTTCGTGGCGGCTTTGCCGTCACTACCGCGGCCCGAGACGT 420
Qy 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db 421 TGGATGCTGCTGGGGGCGCGCTCTCACCGGCTGGCTGCGGTTCCTCCCTCCTAGTG 480
Qy 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
Db 481 GCCCGGFTACATCTCCGAAATCCCTACCCAGCAGTCCGGGGTGTGCTCGCTCCTGT 540
Qy 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
Db 541 GTGAGCTAATGGTCGTCTCGGCACTCTCTGGGCTACTCTGCAGGCTGGGTGCTGGAG 600
Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 601 TGGCGCTGCCTGGCTGTGCTGGGCTCGGTGCCCTCCCTCATGCTGCTCTCATGTGC 660
Qy 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgGlnGluAla-Metal 220
Db 661 TTCATGCCGAGACCCCGCTTCCTGTGACTCAGCACAGGCGCCAGAGGCTGCTC--- 718
Qy 220 aAlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGlu 240
Db 718 ----- 718
Qy 240 uGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleLeu 260
Db 718 ----- 718
Qy 260 yValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
Db 718 ----- 718
Qy 280 uThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300

```

Best Local Similarity: 71.51% Mismatches: 0
Query Match: 51.61% Indels: 103
DB: 13 Gaps: 1

US-09-886-954A-1 (1-477) x US-10-169-395-19 (1-789)

Qy 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db 1 ATGAGCCCGAGGACCCAGAGAAACCCAGCGCTTCCTGGGGCTCTCGGGCGAGCCGC 60

Qy 21 ProArgGlyArgArgValPheLeuAlaAalaPheAlaAalaLeuLeuGlyProLeuSerPhe 40
Db 61 CCGCGCGCCCGCGGTCTTCCTGCGCGCTTCGCGCTTCGCGCTTCGCGCTTCGCGCTTC 120

Qy 41 GlyPheAlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
Db 121 GCTTCGGCTTCGGCTACAGCTCCCGCGCCATCCCTAGCCCTGCAGCGCGCGCGCCCG 180

Qy 61 AlaProArgLeuAspAlaAalaAalaSerTrpPheGlyAlaValValThrLeuGlyAla 80
Db 181 GCGCGCGCTGGAGACCGCGCCCTCTGTTTCGGGGCTTCGTGACCTCGGTGCC 240

Qy 81 AlalaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLeuSerLeu 100
Db 241 GCGCGGGGGAGTGTCTGGCGGCTGGCTGTGGACCGCGCGCGCGCGCAAGCTGAGCCTC 300

Qy 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db 301 TTGCTGTCTCCGTGCGCTTCGTGGCGGCTTCGTGCGGTATCATCAGCGCGCGCCAGGAGCTG 360

Qy 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db 361 TGGATGCTGCTGGGGGCGCTCTCTACCGGCTGGCTGGGTGTTCCTCCTAGTG 420

Qy 141 AlaProValTyrIleSerGluAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
Db 421 GCGCGGTCTACATCTCCGAAATCGCTTACCAGCAGTCCGGGGTGTCTCGGCTCCTGT 480

Qy 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
Db 481 GTGACGCTAATGGTGTCTCGGATCTCTCTGCGCTTACCTGGCAGGCTGGGTGCTGAG 540

Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 541 TGGCGCTGGCTGTGTGGGTGGCTGCGTCCCGCTCCCTCATGCTGTCTCATGTGC 600

Qy 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgGlnGluAla-MetAl 220
Db 601 TTCTATGCCGAGACCCCGCGCTTCTGTGACTCAGCACAGCGCCAGAGGCTGCTC-- 658

Qy 220 aAlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProPheGlyAlaGl 240
Db 658 ----- 658

Qy 240 uGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIysProPheIleGl 260
Db 658 ----- 658

Qy 260 yValSerLeuMetAlaPheGlnGlnLeuSerGlyValAlaValMetPheTyrAlaGl 280
Db 658 ----- 658

Qy 280 uThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVa 300
Db 658 ----- 658

Qy 300 lIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgAlaGl 320
Db 658 ----- 658

Qy 320 uLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPh 340
Db 659 ----CTGCTTGTTCAGGTGGTTCATGGTGTTCAGCAGGAGTGCCTTCGCGCGCTACTT 714

Qy 340 eLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaLeSerAlaProValSe 360
Db 715 CAAGCTGACCCAGGGTGGCCCTCGCACTCCTCGACGTCGTCATCTCGGGCCTGTCTC 774

Qy 360 rAlaGlnProVal 364
Db 775 TGCACAGCTGTT 787

RESULT 7

US-10-157-031-358
; Sequence 358, Application US/10157031
; Publication No. US2003010890A1
; GENERAL INFORMATION:

; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashev, A. V.
; APPLICANT: Krukovskaya, L. L.

; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
; FILE REFERENCE: 2760-103

; CURRENT APPLICATION NUMBER: US/10157,031

; CURRENT FILING DATE: 2002-05-30

; NUMBER OF SEQ ID NOS: 415

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 358

; LENGTH: 2356

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-157-031-358

Alignment Scores:
Pred. No.: 1,01e-92 Length: 2356
Score: 958.50 Matches: 227
Percent Similarity: 59.22% Conservative: 78
Best Local Similarity: 44.08% Mismatches: 164
Query Match: 39.01% Indels: 47
DB: 15 Gaps: 12

US-09-886-954A-1 (1-477) x US-10-157-031-358 (1-2356)

Qy 3 ProGlu---AspProGluGluThrGlnProLeuLeuGly----- 14

Db 12 CCCGAGAGACCCCGCCATGCGAGGACCGCTGTGGAGCCGAGGCGCGGACCTGCAG 71

Qy 15 -----ProProGlyGlySerAlaProArgGlyArg----- 24

Db 72 ACCTTCCCGAGAAAGCCGCCCGCTGCCAGGGGACAGGCGCGGTCGGACCTGCAG 131

Qy 25 -----ArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPhe 42

Db 132 AACAAAAGGTGTTCCTGGCCACCTTCGCGCAGTGTCTCGCAATTTTCAGCTTTGGGTAT 191

Qy 43 AlalaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProAlaPro 62

Db 192 GCCTGTGTCTACATCCCTGTCAUCCAGCCCTGGAGCGCTCCTTGGATTCCTGACCTG 251

Qy 63 ArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAla 82

Db 252 CATCTGACCAAAATCCAGGCATCCTGTGTTGGTCCGTGTTCACCTGGGAGCAGCGGCC 311

Qy 83 GlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgIysLeuSerLeuLeu 102

Db 312 GGAGCGCTGAGTGCATGATCCTCAACGACCTCTGGGCGGAGGCTGAGCATCATGTT 371

Qy 103 CysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122

Db 372 TCAGCTGTGCTGCGCGCGCTGATGCGCTCATGCGCGGTGCGCACGCGCTCTGGATG 431

Qy 123 LeuLeuGlyGlyArgLeuLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaPro 142

Db 432 CTGCTGTTCGAAAGGAGCTGACGGGCTTCGCGGGGGGTCTACAGCTGCTGATCCCG 491


```
Db 500 GTGTACGTGTCTGAGATTGCTCCCGCAGCGCTCTGTGGGGCTCTGGGGCCACACCCCGAG 559
Qy 163 LeuMetValValGlyLeuLeuAlaTyrLeuAlaGlyTrpValLeuLeuTrpArg 182
Db 560 CTCATGGCAGTGTTCGATCCCTGTCCCTACGCGCTTGGCCCTCTCTGCCGTGGGCG 619
Qy 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
Db 620 TGGCTGGCTGTGGCGGGAGCGCTGTGTCTATCATGATCCCTGTCTCTCAGCTTCATG 679
Qy 203 ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaAlaLeu 222
Db 680 CCCAACTCGCGCGCTTCTCTCTCGGGCAGGACGAAGAGCCCTCGGGCGCTG 739
Qy 223 ArgPheLeuTrpGlySerGluGlnGly-----TrpGluAspProProIleGly----- 238
Db 740 GCCTGTGGCTGTGGCAGCGAGCTGCATGTCACCTGGGAGTTTCGAGCAGATCCAGACAAC 799
Qy 239 --AlaGluGlnSerPheHisLeu-----AlaLeuLeuArgGlnProGlyIleTyrLys 255
Db 800 GTCCGGACAGACGACCGAGTATCTGGGCTGAGGCACGGGCCACACAGTGTGCCGG 859
Qy 256 ProPheIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaVal 275
Db 860 CCCATCACCGTGGCTTCTGTGCTCGCTCGCAGCAGCTGACGGGCATCACGCCCATC 919
Qy 276 MetPheTyrAlaGluThrIlePheGluGluAlaLysPhe-----LysAspSer 291
Db 920 CTGTGTACCTGCAGTCCATCTTCGACAGCACCGCTGTCCTGCTGCCCGCCCAAGCAGC 979
Qy 292 SerLeuAlaSerValValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeu 311
Db 980 -----GCAGCATCGTGGGCGCTGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1030
Qy 312 IleMetAspArgAlaGlyArgLeuLeuLeuValLeuSerGlyValValMetValPhe 331
Db 1031 ACCATGGACCTCGCAGCGCGCAAGTGTCTCTTCGCTCAGCGGCCATCATGTTTGTCT 1090
Qy 332 SerThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyGlyProGlyAsnSerSer 351
Db 1091 GCCAACCTGACTCTGGGGCTGTATCATCCATTT-----GGCCCCAGCGCCTCTGAGC 1141
Qy 352 HisValAlaIleSerAlaProValSer-----AlaGlnProValAsp 365
Db 1142 CCCACAGCATCTGGGCGCTGGAAGCAGTCTCTGGGGGACTGTGGCGCAGCCCTTGCA 1201
Qy 366 AlaSerValGly---LeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAlaGly 384
Db 1202 GCACCGGTGGCTACCTCACCTGTGTGCCCTGTGCTGGCCACCATGCTCTTCATCATGGC 1261
Qy 385 PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePheProLeuHis 404
Db 1262 TACCGCGGGCTGGGTCCTCATCCCTGGTGTCTCATGTCGTGAGGTCCTCCCGCTGCTG 1321
Qy 405 ValLysGlyValAlaThrGlyLeuCysValLeuThrAsnTrpLeuMetAlaPheLeuVal 424
Db 1322 GCCGTGGCGTGGCTCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1381
Qy 425 ThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeuAla 444
Db 1382 ACCAAGTCTTCTCTGCCATGGT--GAGCACCTTCGGCTCCAGGTGCTCTTCTTCTTCT 1440
Qy 445 -SerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrLysG 464
Db 1441 GCGCGCATCTGCTGTGGTGGCTGTGTTTACAGGCTGCTGTGTGCTGCTGCTGCTGCTG 1500
Qy 464 YLysThrLeuGluGlnIleThrAlaHisPheGlu---GlyArg 477
Db 1501 ACGGTCTCTGGAGCAGATCGAGTCTCTTTCGTCATGGGGAGA 1543

RESULT 9
US-10-172-118-1685
; Sequence 1685, Application US/10172118
```

```
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1685
; LENGTH: 2487
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_017585
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1685

Alignment Scores:
Pred. No.: 1,09e-92 Length: 2487
Score: 958.50 Matches: 227
Percent Similarity: 59.22% Conservative: 78
Best Local Similarity: 44.08% Mismatches: 164
Query Match: 39.01% Indels: 47
DB: 13 Gaps: 12

US-09-886-954A-1 (1-477) x US-10-172-118-1685 (1-2487)
Qy 3 ProGlu---AspProGluGluThrGlnProLeuLeuGly----- 14
Db 20 CCCGAGAGACCCCGGCCATGTCAGGAGCGCTGTCTGGAGCCGAGGGCCCGGACTACGAC 79
Qy 15 -----ProProGlyGlySerAlaProArgGlyArg----- 24
Db 80 ACCTTCCCGGAGAGCCGCCCTCCGCCAGGAGCGCGGTCGGACCCCTGCAG 139
Qy 25 -----ArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPhe 42
Db 140 AACAAAAGGGTGTTCCTGGCCACCTTCGCCGACGTCTCGCAATTTTCAGCTTTGGGTAT 199
Qy 43 AlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaPro 62
Db 200 GCCCTGTCTACATCCCTGTCTATCCAGCCCTGGAGCGCTCTTGGATCCTGACCTG 259
Qy 63 ArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyValAlaAla 82
Db 260 CATCTGACCAATCCAGGCATCTCTGTTGGTTCGCTGTTCACCTGGGAGAGCGGCC 319
Qy 83 GlyGlyValLeuGlyGlyTrpLeuValAspAlaGlyArgLysLeuSerLeuLeuLeu 102
Db 320 GGAGCGCTGATGTCATGATCTCAAGCACCTCTCGGGCCGGAAGCTCAGCATCATGTC 379
Qy 103 CysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
Db 380 TCAGCTGTGCTGCGCGCGGCTATGCGCTCATGGCGGTGCGCAGCGGCTCTGGATG 439
Qy 123 LeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaAlaSerLeuValAla 142
Db 440 CTGCTGCTCGAAGAGCAGTCAGCGGCTTCGCCGGGGGCTCACAGCTGCCTGCATCCCG 499
Qy 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
Db 500 GTGTACGTGTCTGAGATTGCTCCCGCAGCGGCTCTGTGGGGCTCTGGGGGCCACACCCAG 559
Qy 163 LeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuLeuTrpArg 182
```



```
QY 83 GlyGlyValLeuGlyGlyTrpLeuValAlaAspArgAlaGlyArgLysLeuSerLeuLeuLeu 102
Db 355 GGAGGCGCTGAGCGCATGATCTCAACAGCCTCCCTGGCGCGAAGCTGAGCATGATGTC 414
QY 103 CysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
Db 415 TCAGCTGTGCCGTGGCGGCGCGCTATGGCTCATGGCGGGTGGCGACGGCTCTGGATG 474
QY 123 LeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaPro 142
Db 475 CTGCTGCTCGAAGACGCTGACGGGCTTCGCGCGGGGCTCACAGCTGCTGCATCCCG 534
QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
Db 535 GTGTAGCTGTCTGAGATGCTCCCGCAGCGCTTCGTGGGGCTCTGGGGCCACACCCAG 594
QY 163 LeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArg 182
Db 595 CTCATGGCAGTGTGGATCCCTGTCCCTACGGCCTTGGCTCTGCTGCGGTGGCG 654
QY 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
Db 655 TGGCTGGCTGTGGCGGAGCGCGCTGTGCTCATCATGATCTGCTGCTCAGCTTCATG 714
QY 203 ProGluThrProArgPheLeuLeuThrGlnHisArgGlnGlnAlaMetAlaAlaLeu 222
Db 715 CCAACTCGCGCGCTCTCTGCTCTCGGGCAGGACAGAGGCCCTGGGGCGCTG 774
QY 223 ArgPheLeuTrpGlySerGluGlnGly-----TrpGluAspProProlGly----- 238
Db 775 GCCTGCTCGTGGAGCGAGCTGATGTCACCTGGAGTTCAGCAGATCAGGACAAC 834
QY 239 ---AlaGluGlnSerPheHisLeu-----AlaLeuLeuArgGlnProGlyIleTyrLys 255
Db 835 GTCCGAGACAGACGCGCAGTATCGTGGCTGAGGCACGGGCCGCCACACGTGTCGG 894
QY 256 ProPheIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaVal 275
Db 895 CCATCACCCTGCTGCTGTGAGCGCCCTCTGACAGCTGACGGGCATCAGCCCATC 954
QY 276 MetPheTyrAlaGluThrIlePheGluGluAlaLysPhe-----LysAspSer 291
Db 955 CTGGTCTACCTGACGTCATCTTCGACAGCAGCGCTGCTGTCGCCCGCCCGACGAC 1014
QY 292 SerLeuAlaSerValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeu 311
Db 1015 -----GCAGCCATCGTTGGGCGCGTGGCGCTCTGTCGCTGATCGCGCCCTC 1065
QY 312 IleMetAspArgAlaGlyArgArgLeuLeuValLeuSerGlyValValMetValPhe 331
Db 1066 ACCATGGACTCGCAGCGCGCAGGCTGCTCTTCGCTCAGCGGCCATCATGTTGCT 1125
QY 332 SerThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSer 351
Db 1126 GCCAACCTGACTCTGGGGTGTATCATCCACTTT-----GGCCCCAGGCGCTGAGC 1176
QY 352 HisValAlaIleSerAlaProValSer-----AlaGlnProValAsp 365
Db 1177 CCCAACAGACTCGCGGCGCTGGAAGCAGAGTCTCGGGGGAGCTTGGCGCAGCGCCCTG 1236
QY 366 AlaSerValGly---LeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAlaGly 384
Db 1237 GCACCGCTGGCTACCTCACCCTGGTGGCCCTGCTGCGCCACCATGCTCTTCATCATGG 1296
QY 385 PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePheProLeuHis 404
Db 1297 TAGCGCTGGCTGGGGTCCCATCATCGCTGCTCATGCTCTGAGGCTCTGCGCCCTGCGT 1356
QY 405 ValIysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeuVal 424
Db 1357 GCCCGTGGCGTGGCGCTCTGCTGGCTGCTGGCCAGCTGGCTCACCGCTTCGCTCTC 1416
```

```
QY 425 ThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeuAla 444
Db 1417 ACCAAGTCTTCTCCAGTGGT-CAGCACCTTCGGCTCCAGGTCGCTTCTTCTTCT 1475
QY 445 -SerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrLysGln 464
Db 1476 CGCGGCCATCTGCTTGGTGAGCGCTGCTTACAGGCTGCTGTGTGCCCGACGACCAAGG 1535
QY 464 LysThrLeuGluGlnIleThrAlaHisPheGlu---GlyArg 477
Db 1536 ACGGTCCTGGAGCAGATCGAGTCTTCTTCGCGCGGGGAGA 1578
```

RESULT 12

```
US-10-424-599-132422
; Sequence 132422, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 132422
; LENGTH: 2202
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2202)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90584C.1
US-10-424-599-132422
```

Alignment Scores:

Pred. No.:	1,2e-66	Length:	2202
Score:	716.00	Matches:	164
Percent Similarity:	54.88%	Conservative:	89
Best Local Similarity:	35.57%	Mismatches:	174
Query Match:	29,14%	Indels:	34
DB:	13	Gaps:	7

US-09-886-954A-1 (1-477) x US-10-424-599-132422 (1-2202)

```
QY 26 ValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGly 45
Db 384 GTCTCTTCTGCTGCTCATCGCTGGCTGGGTGCCATTCAATTCGGCTTCAGGTGGG 443
QY 46 TyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArgLeuAsp 65
Db 444 TATCTCTCCAAACCAAGGGCTATAGTTCGAGTCTA-----AACCTCTCT 491
QY 66 AspAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAlaGlyGlyVal 85
Db 492 ATTTCGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 551
QY 86 LeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuLeuCysSerVal 105
Db 552 GCTAGTGGTCAGATGATACATACATCGCGCGCAAGGGTCATTGATGATGCTGCGATC 611
QY 106 ProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGly 125
Db 612 CCCAATATAATAGGGTGGCTTGTCTATTTCTTTTTCGCAAGATTCCTCGTTTGTATATG 671
QY 126 GlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaProValTyrIle 145
Db 672 GGAGGTTGTTGAGAGTTTTCGGCTGGGATTTATCTTATGTTGGTGGCTGTTTATATA 731
```

```

Qy 146 SerGluIleAlaTyProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetVal 165
Db 732 GCTGAGATTGACCTCAAACTTGAGAGTGGCCCTGGGTGAGTGAACACAGCTCTCTGTT 791
Qy 166 ValValGlyIleLeuLeuAlaTyTrpValLeuGluTrpValArgTrpLeuAla 185
Db 792 ACAATTGGCATTATGCTGGCTTATCTGTGGGTCTTTTGTCAACTGGAGAGTCTTGCA 851
Qy 186 ValLeuGlyCysValProSerLeuMetLeuLeuMetCysPheMetProGluThr 205
Db 852 AITCTAGAAATTTGGCCTTGACATTAATACCTGGATTATTTTTCATACCTGAATCC 911
Qy 206 ProArgPheLeuLeuThrGlnHisArgGlnGluAlaMetAlaAlaLeuArgPheLeu 225
Db 912 CCCAGATGGTGGCAAGATGGGATGATAGATGAGTTTGACATCTCTTTCAGATGTTA 971
Qy 226 TrpGlySerGlnGlyTrpGluAspProPhe 237
Db 972 CGAGGATTGAC-----ACTGATATATCTGTGGAAGTACATGAATAAAGAGATCT 1022
Qy 238 -----GlyAlaGluGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIle 253
Db 1023 GTGGCTTCAACCGGAAAAAGAGCTCCAAATCCGATTTCAGATCTCAAGAGGAAAAAGATAT 1082
Qy 254 TyrLysProPheIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsn 273
Db 1083 TGGTTCGGTTAAATGGTGGTATGGATTACTGTCTTCAGCAATTAATCTGGTATCAAT 1142
Qy 274 AlaValMetPheTyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerLeu 293
Db 1143 GGAATTTTGTCTATTCAACTACCATCTTTGCAATGACGAAATTTTCATCCAGCGAAGCT 1202
Qy 294 AlaSerValValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuMet 313
Db 1203 GCTACAGTTGAGCTGGAGCGCTCAGCTCATAGCACTGGAATTTCCACATGGTGTG 1262
Qy 314 AspArgAlaGlyArgArgLeuLeuValLeuSerGlyValValMetValPheSerThr 333
Db 1263 GACAAAGTGGCGGAGGCTGCTCTATATATCTCTATATATCTCTATATGACAGTTAGCCTT 1322
Qy 334 SerAlaPheGlyAlaTyPheLysLeuThrGlnGlyGlyProGlyAsnSerSerHisVal 353
Db 1323 CTCATTGTTTCTATAGCATTTTATCTG--GAGGGGGTGTATCATCAGAGGATTCACATTTA 1379
Qy 354 AlaIleSerAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeu 373
Db 1380 -----TTCAGCATTTTGGGAATAGTTTCTATTGTGTGGACTCGTGCTATG 1424
Qy 374 AlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValAlaGlyTrpGlyProIlePro 393
Db 1425 GTGATT-----GGGTCTCTCTAGTCTGGGACCCATCCCT 1460
Qy 394 TrpLeuLeuMetSerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCys 413
Db 1461 TGGCTTATAATGCTGAGATACTTCCAGTGAATATAAAGGCGCTTGCTGGCAGCATAGCG 1520
Qy 414 ValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGlu 433
Db 1521 ACAATGGGAATTTGGCTGATTCGGGGGATTCAGATGCTGCTAACTTCTGTTGAAT 1580
Qy 434 ValLeuArgProTyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu 453
Db 1581 ---TGGAGCAGTGGAGGACATTTACATCTACAGCTGCTAGTGGCTTACTATAGCT 1637
Qy 454 PheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHis 473
Db 1638 TTTATAGCAATGTGGGTTCTCTGAGACCAAGGAGAACATTGGAGAAATTCAGTTTCC 1697
Qy 474 Phe 474
Db 1698 TTC 1700

```

RESULT 13

```

US-10-425-114-26596
; Sequence 26596, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26596
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4371-017-Fl1_FLI
US-10-425-114-26596

```

```

Alignment Scores:
Pred. No.: 2,9e-65 Length: 1797
Score: 702.00 Matches: 170
Percent Similarity: 53.56% Conservative: 78
Best Local Similarity: 36.72% Mismatches: 176
Query Match: 28.57% Indels: 40
DB: 13 Gaps: 6

```

US-09-886-954A-1 (1-477) x US-10-425-114-26596 (1-1797)

```

Qy 27 PheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGlyTyr 46
Db 284 TTCCTCTGCAGCTCATCGTCGCGCTGGCCCATCAGTTCGATTCACGGGGGCTTC 343
Qy 47 SerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArg-LeuAspAs 66
Db 344 TCCTCCCGACG-----CAGGACGCCATCAT-CCGAGACCTTCGACCTCACCT 390
Qy 66 pAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAlaGlyGlyValLe 86
Db 391 CTCGAGTTCGTGTGCGATCGCTGTCACCGCGCCCATGTTGGGCGCATTC 450
Qy 86 uGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuLeuCysSerValPr 106
Db 451 CAGTGGTCAGATGCCGAGTACATTTGGGCGCAAGGGTCATTGATGATGTCGCAATCC 510
Qy 106 pPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGlyGl 126
Db 511 AAACATCATTTGGTGGCTTGCCATCTCTTTGCAAAAGACTCATCGTTCTTTATATGGG 570
Qy 126 ValGlyLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaProValTyrIleSe 146
Db 571 ACGATTGCTCAGGGGTTTGGTGTGGTGTCTCTATATACCGTGCAGTTTACATAGC 630
Qy 146 rGluIleAlaTyProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetValVa 166
Db 631 AGAATATACCTCAAAACATCAGAGGTGCTCTTTGGCTCAGTGAATCAGTTATCTGTAA 690
Qy 166 lValGlyIleLeuLeuAlaTyTrpValLeuGluTrpValLeuGluTrpArgTrpLeuAlaVa 186
Db 691 CGTTGGTATATTTGTCGATATTTGCTCGGCATGTTTGTCTCTGGAGGCTCTCTGCTGT 750
Qy 186 lLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMetProGluThrPr 206
Db 751 AATAGGAATCTTGCCTTGCATGTGTGTGATACCTTCGGCTTATCTTCATTCAGAAATCCC 810
Qy 206 oArgPheLeuLeuThrGlnHisArgGlnGlnAlaMetAlaAlaLeuArgPheLeuTr 226
Db 811 AAGATGGTTGGCAAGATGAACATGATGATGATTTTGAGACTTCTTTACAAAGTTCGAG 870

```

```
QY 226 pGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu----- 240
DB 871 GGGATTGAG-----ACTGACATACGCGGGAAGTGAATGATATAAAGAG 915
QY 241 -----GlnSerPheHisLeuAlaLeuLeuLeuArgGlnProGI 252
DB 916 AGCAGTAGCGTCAGCAAAACAAAGGACACGATCGCTTTTCAAGAAATTAACACCAAGAA 975
QY 252 yIleTyrLysProPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyVa 272
DB 976 ATACCGCACACCCCTAATACTAGGAATTGGCTACTTGTACTGCAACAGCTAAAGTGAAT 1035
QY 272 tAsnAlaValMetPheTyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSe 292
DB 1036 CAATGGAAATATGTTTATGAGTAGCATCTTCAAGCAGCAGCTCTCACAAACAGTGA 1095
QY 292 rIeuAlaSerValValValGlyValIleGlnValLeuPheThrAlaValAlaLeuI 312
DB 1096 CTTGGCTACATGTGCACCTGGTGTATCCAGGTTCTTGTACAGGAGTTTACAACCTGGTT 1155
QY 312 eMetAspArgAlaGlyArgArgLeuLeuValLeuSerGlyValValMetValPheSe 332
DB 1156 ATTACACAGAGCTGCGCAGCAGATCTCTTATCATCTCTTCTGTGGATGACTCTAAG 1215
QY 332 rThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyGlyProGlyAsnSerSerHi 352
DB 1216 CTTCTTGCAGTGTGCTGTGTATTTTCTCTC---AAGGATAGCATTTCAAGATTTCTCA 1272
QY 352 sValAlaIleSerAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTr 372
DB 1273 CATGTACTACACC-----TTAAGTAT 1293
QY 372 pLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValAlaGlyTrpGlyProII 392
DB 1294 GATCTCCTTGGTGTCTTGTGGCTTTGTAAATCGCCTTCTCCTCGGTATGGGTCACAT 1353
QY 392 eProTrpLeuLeuMetSerGluIlePheProLeuHisValLysGlyValAlaThrGlyI 412
DB 1354 TCCATGGATCAATAGTCAGAGATCTCCGGTGTAGTATCAGAGCTCTCGCAGGAGCTT 1413
QY 412 eCysValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMe 432
DB 1414 TCGCAGCGCTCGCCAACTGGCTTACATCTCTTGGAAATAACAATGACA--GCAAACTTGAT 1470
QY 432 tGluValLeuArgProTyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerVa 452
DB 1471 GCTTAGCTGGAGTGTGGGGGACCTTTGTGTCTACATGGTGTGAGTGTCTTCCACCT 1530
QY 452 lLeuPheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThrAl 472
DB 1531 CGTGTTCGTATCTCTTGGGTGCCAGACAAAGGAAGAACTCTCGAAGAGATACAATG 1590
QY 472 ahisPhe 474
DB 1591 GTCCTTC 1597
```

RESULT 14

```
US-10-437-963-61061
; Sequence 61061, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
```

```
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 61061
; LENGTH: 2003
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_62530C.1
US-10-437-963-61061
```

Alignment Scores:

```
Pred. No.: 3,38e-65 Length: 2003
Score: 702.00 Matches: 170
Percent Similarity: 53.56% Conservative: 78
Best Local Similarity: 36.72% Mismatches: 176
Query Match: 28.57% Indels: 40
DB: 17 Gaps: 6
```

US-09-886-954A-1 (1-477) x US-10-437-963-61061 (1-2003)

```
QY 27 PheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGlyTyr 46
DB 346 TTCTCTGCACGCTCATCGTCGCTCGGCCCATCCAGTTCCGATTCCACGGGGGGCTTC 405
QY 47 SerSerProAlaIleProSerLeuGlnArgAlaAlaProAlaProArg-LeuAspAs 66
DB 406 TCTTCCCGCAGC-----CAGACGCGCATCAT-CCGAGACCTTCGACCTCACCGT 452
QY 66 pAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAlaGlyGlyVal 86
DB 453 CTCGAGGTTCTCGGTGTTCCGATCGCTCCAACTCGCGGCCCATGGTTGGGGCGATTGC 512
QY 86 uGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuLeuCysSerValPr 106
DB 513 CAGTGTGCAGATGGCGCGGAGTACATTGGCGCGCAAGGGTCATTGATTGCTGCAATTC 572
QY 106 oPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGly 126
DB 573 AAACATCATTTGGTGGCTTGCATCTCTTTGCCAAAGACTCATCGTTCTTTATATGGG 632
QY 126 yArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaProValTyrIleSe 146
DB 633 ACGATTGCTCGAGGGGTTGGTGTGGTGTCTCTTATACGGGTGCCAGTTTACATAGC 692
QY 146 rGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetValVa 166
DB 693 AGAAATATCACCCTCAAAACATGAGAGGGTCTTTGGCTCAGTGAATCAGTTATCTGTAA 752
QY 166 lValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArgTrpLeuAlaVa 186
DB 753 CGTTGGTATATTGTTGGCATATTGCTCGGCATGTTTGTCTTCCGTGAGGCTTCTTGCTGT 812
QY 186 lLeuGlyCysValProProSerLeuMetLeuLeuLeuMetCysPheMetProGluThrPr 206
DB 813 AATAGGAATCTTGGCTTGCATCTGTGTTGATACCTGGCCCTATTCTTCATTCAGAAATCCC 872
QY 206 oArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuTr 226
DB 873 AAGATGGTTGGCAAGATGAACATGATGGATGATTTTGAGACTTCTTTACAGTTCTTGAG 932
QY 226 pGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu----- 240
DB 933 GGGATTTCAG-----ACTGACATCAGCGCGGAGTGAATGATATATAAAGAG 977
QY 241 -----GlnSerPheHisLeuAlaLeuLeuLeuLeuArgGlnProGI 252
DB 978 AGCAGTAGCGTCAGCAAAACAAAGGACACATCGCTTTTCAAGAAATTAACACCAAGAA 1037
QY 252 yIleTyrLysProPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyVa 272
DB 1038 ATACCGCACACCCCTAATACTAGGAATTGGCTACTTGTACTGCAACAGCTAAAGTGAAT 1097
```



```
QY 309 AlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuValLeuSerGlyValVal 328
Db 1383 ACAACCTGGTCTGACGACAAAGCTGTCGAGGCTCTCTCATTTATTTCCACCACAGA 1442
QY 329 MetValPheSerThrSerAlaPheGlyAlaTyPheLysLeuThrGlnGlyGlyProGly 348
Db 1443 ATGGTCATTACTCTTGTTATTCTTCTGTGTCATTTTGTGAAGGACACACATAGCTGCT 1502
QY 349 AsnSerSerHisValAlaIleSerAlaProValSerAlaGlnProValAspSerVal 368
Db 1503 GGTTCGCACCTTATCTCTGTA----- 1523
QY 369 GlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGly 388
Db 1524 ---ATGAGTAGCTTTTCACTGGCTGGACTGTGGCATTTGTGATTCATTTCTCTGGC 1580
QY 389 TrpGlyProIleProTrpLeuLeuMetSerGluIlePheProLeuHisValLysGlyVal 408
Db 1581 TTGGGAGCGATTCCGCTGGATCATTATGCTGAGATCCTCTCTGTTAATCAAGAGCCTT 1640
QY 409 AlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPhe 428
Db 1641 GCTGAAGTGTTCGCGACCTGCGCACTGGCTGCAGCATGGCCCATTAATGACG--- 1697
QY 429 SerSerLeuMetGluValLeuArgProTyrglyAlaPheTrpLeuAlaSerAlaPheCys 448
Db 1698 GCAAGCCTGATGTTGAACTGGAGCAGCTGGAGGAACATTTGCTATCTACGCCGCTGTGCT 1757
QY 449 IlePheSerValLeuPheThrLeuPheCysValProGluThrLysGlyLysThrLeuGlu 468
Db 1758 ACCATGGCCCTCATTTTCTGTGCTTGTGGTGGCTGAGACCAAGGAAGAACGCTAGAG 1817
QY 469 GlnIleThrAlaHisPhe 474
Db 1818 GAAATCGCCTTCTCATTC 1835
```

Search completed: September 28, 2004, 08:02:21
Job time : 4267 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 28, 2004, 01:47:14 ; Search time 2846 Seconds
(without alignments)
5005.011 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPEDPRTQPLGPPGSA.....CVPETKKTLEQITAHFEGR 477

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlp
-Q=/cgn2_1/USPIO_spool_pUS09886954/runat_27092004_170632_13400/app_query.fasta_1.647
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09886954.qcgn1_5180_runat_27092004_170632_13400 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2159	87.9	2101	11	AK081806	AK081806 Mus muscu
2	2123.5	86.4	1954	11	AF289587	AF289587 Homo sapi
3	1713	69.7	1008	29	AY414180	AY414180 Homo sapi
4	1490.5	60.7	1201	13	BX395379	BX395379 BX395379
5	1485.5	60.5	1000	29	AY414182	AY414182 Mus muscu
6	1379	56.1	1085	12	BM545247	BM545247 AGENCOURT
7	1340	54.5	932	12	BG479842	BG479842 602527373
8	1293	52.6	830	12	BG700749	BG700749 602681616
9	1269.5	51.7	991	12	BI334832	BI334832 602998959
10	1269	51.6	1004	29	AY414181	AY414181 Pan trogl
11	1267	51.6	875	14	CD557801	CD557801 AGENCOURT
12	1241.5	50.5	880	12	BM044230	BM044230 603621486
13	1238	50.4	828	12	BG478000	BG478000 60252606
14	1233	50.2	777	12	BG717034	BG717034 602688093
15	1230.5	50.1	889	12	BI757409	BI757409 603029344
16	1222.5	49.8	879	12	BG771736	BG771736 602720404
17	1193.5	48.6	838	12	BG281777	BG281777 602402292
18	1183.5	48.2	801	12	BI114765	BI114765 602861314
19	1120.5	45.6	894	13	BQ919165	BQ919165 AGENCOURT
20	1119	45.5	899	12	BI916408	BI916408 603178364
21	1116	45.4	854	12	BI828656	BI828656 603078566
22	1076	43.8	1002	12	BI831908	BI831908 603078776
23	1074.5	43.7	1002	12	BI828949	BI828949 603074819
24	1059	43.1	833	12	BI826390	BI826390 603076257
25	1055	42.9	700	13	BQ109572	BQ109572 imagegc 7
26	1036.5	42.2	934	10	BE910478	BE910478 601501005
27	1030.5	41.9	773	13	BQ603775	BQ603775 MI-P-CP1-
28	990	40.3	664	10	BE986058	BE986058 UI-M-CG0P
29	980	39.9	621	10	BE986074	BE986074 UI-M-CG0P
30	973	39.6	920	10	BF308306	BF308306 601890370
31	949	38.6	2125	11	AK089246	AK089246 Mus muscu
32	948	38.6	677	12	BG707107	BG707107 602670234
33	947.5	38.6	684	12	BI907581	BI907581 603065948
34	938	38.2	735	12	BM008943	BM008943 603618748
35	933.5	38.0	2081	11	AK079650	AK079650 Mus muscu
36	931	37.9	994	13	BQ644799	BQ644799 AGENCOURT
37	912.5	37.1	683	13	BY750238	BY750238 BY750238
38	900	36.6	638	10	BB612439	BB612439 BB612439
39	889	36.2	608	10	BG079217	BG079217 H3037003-
40	888.5	36.2	908	14	CA327128	CA327128 UI-M-FY0-
41	870	35.4	974	13	BQ944156	BQ944156 AGENCOURT
42	869	35.4	594	12	BI402088	BI402088 MI-P-CP0-
43	839	34.1	917	10	BF140667	BF140667 601786917
44	837.5	34.1	584	14	CB586069	CB586069 AMGNNUC:N
45	834.5	34.0	647	14	CA322312	CA322312 UI-M-FX0-

ALIGNMENTS

RESULT 1
AK081806
LOCUS
DEFINITION Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone: C130078K14 product: solute carrier family 2, (facilitated glucose transporter), member 8, full insert sequence.
ACCESSION AK081806
VERSION AK081806.1 GI:26349396
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 JOURNAL MEDLINE
 99279253
 PUBMED
 10349636

REFERENCE

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL MEDLINE
 20493374
 PUBMED
 11042159

REFERENCE

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 JOURNAL MEDLINE
 20530913
 PUBMED
 11076861

REFERENCE

4 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 JOURNAL MEDLINE
 20530913
 PUBMED
 11076861

REFERENCE

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 JOURNAL MEDLINE
 20530913
 PUBMED
 11076861

REFERENCE

6 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozawa, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasumishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)

REFERENCE

7 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 Direct Submission
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)

COMMENT

CDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/

FEATURES

Source

1. 2101
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:C130078K14"

/db_xref="MGI:2414880"
 /db_xref="taxon:10090"
 /clone="C130078K14"
 /tissue_type="head"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="16 days embryo"
 61. 1494
 /note="unnamed protein product; putative
 solute carrier family 2, (facilitated glucose
 transporter), member 8 (MGI:MGI:1860103, GB|NM_019488,
 evidence: BLASTN, 99%, match=1490)"
 /codon_start=1
 /protein_id="BAC38338.1"
 /db_xref="GI:26349397"
 /translation="MSPEDQETOPLEPPPEARTPRGRVFLASFAAALGPLSFGFAL
 GYSPALPSLRTAPPALRLGDNAASWFGAVVTIGAAAGGLGWLDRAGKLSULL
 CTVPVTFGAVITAAQDVMMLGRLLTGLACGVASLVAPVYISIAIPAVRGLGSC
 VQMVVITGLAYVAGWLEWRLAVLGVPTMLLMCMYPTETPLTTHQVQEA
 MAALRFMGSEGEWEPVAGHQFQVALLRPGLIKPLIGISLMVFOQLSGVNAI
 MYANSIFERAKFDSIASVTGIIQVLFATAALIMDRAGRLLLALSIVWFMS
 SARGVYFKLTQSLPSNSHVGIVPIAEPVDVQGLAVGLMCLFTAGPAVGVGPI
 PWLLMSEIFPLHVKGVATGICVLINWFMFLVTKFESSVMEMLRFGAFWLTAAFCAL
 SVLFTLVVPTKRTLEQVTAHFEGR"
 2080..2085
 /note="putative"
 2101
 /note="putative"

polyA_signal

2080..2085

polyA_site

2101

ORIGIN

Alignment Scores:

Pred. No.: 3.49e-185 Length: 2101
 Score: 2159.00 Matches: 412
 Percent Similarity: 92.47% Conservatives: 30
 Best Local Similarity: 86.19% Mismatches: 34
 Query Match: 87.87% Indels: 2
 DB: 11 Gaps: 2

US-09-886-954A-1 (1-477) x AK081806 (1-2101)

QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProGlyGlySerAla 20
 Db 61 ATGTCTCCCGAGAGACCCCGAGAGAGCGCGCTATTTCGGCCACCGGAGACCGAGACT 120
 QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 Db 121 CCCCAGGCGCGCGCGGTCTCTCTCGCTCTCTCGCGCGCGCTCTGGACCCCTCAGCTTC 180
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 Db 181 GGCTTCGGCGCTCGGCTCAGACTCCCGGCATCCCGAGCTCGCGCGCACCGCACCCCG 240
 QY 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 Db 241 GCCCTCGCGCTCGGAGACAATCGGCTCTCTGGTTCGGGCGCGCTCGTACCTCGGCGCT 300
 QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 Db 301 GTGTCAGGGGGCATACTGGCGCGCTGGCTCTCGGCGCTCGAGCGGCGGCGAGCTGAGCTC 360
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 Db 361 TTGCTCTGCACCGCTCGCTTCGTGCTGCTGCTGCTGCTCATCCGCGCGCGCGGATG 420
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 Db 421 TGGATGCTGCTCGAGGGCGCGCTCTCTCACCGCGCTGCTCGGAGTGCCTCCTACTAGTG 480
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 Db 481 GCACCGGTTTACATCTCGGAATCGCTACCCAGCTGTCGAGGAGTGTCTCGCTCCTGT 540
 QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180

```
Db      541  GTGCAGCTGATGGTTGTCTCACTGGCATCTCTCTGCGCTATGTGGCAGGCTGGTCTCTAGAG 600
Qy      181  TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db      601  TGGCGCTGGCTGGCGGTGTGGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 660
Qy      201  PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
Db      661  TACATGCCGAGAGACCCACAGTTTCTCTCTCACTCAACACAGTACACAGAGGCGCATGGCT 720
Qy      221  AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
Db      721  GCCTTGGCTTCTGTGGGCTCTGAGAGGCTGGGAGAGCCCTGTGTGGGCTGAG 780
Qy      241  ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIle 259
Db      781  CACCAGGCTTCCAGCTGGCCCTCTGAGGCGCCTGGCATCTCAAGCCCTCATCATC 840
Qy      260  GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla 279
Db      841  GGCATTTCCTTCATGGTCTCCAGAGCTGTCAAGGCTCAATGCTATCATGTTCTATGCC 900
Qy      280  GluThrIlePheGluAlaValPheLysAspSerSerLeuAlaSerValValValGly 299
Db      901  AACAGCATCTTCGAGAGGCCCAAGTCAAGCAGCAGCAGCCTGGCTCGGTCACTGTGGGC 960
Qy      300  ValIleGlnValLeuPheThrAlaValAlaAlaIleLeuMetAspArgAlaGlyArgArg 319
Db      961  ATAATCCAGGTCCTGTTCACTGCTGTGGCGCCCTCATCATGTGAGCAGAGGCGAAGG 1020
Qy      320  LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr 339
Db      1021  CTGCTCTGGGCTTGTGGGTTGTATCATGTTTCAGTATGAGTGCCCTTTGGTACCTAC 1080
Qy      340  PheLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProVal 359
Db      1081  TTCAAACTGACCCAGAGGCTCTCCAGCAACTCTCCCAAGTAGGCTG---GTGCCATC 1137
Qy      360  SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCys 379
Db      1138  GCGGCGAGCGCTGGATGTCAAAGTGGAGCTGGCTGGCTGGCTGGTGGAGCATGTGC 1197
Qy      380  LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGlu 399
Db      1198  CTCTTCATGCTGCTTTGGCGTGGCTGGGAGCCATCCCTGGCTCTCATGTGAGAG 1257
Qy      400  IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeu 419
Db      1258  ATCTTCCCTCTGCAATGTCAAGGGTGTGGCTACCGGCATCTGTGTCCTCAACCACTGGTTC 1317
Qy      420  MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGly 439
Db      1318  ATGGCTTCTTAGTGACCAAGAGTTTCAGAGCGCTCATGGAGATGCTCAGACCCCTACGGT 1377
Qy      440  AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
Db      1378  GCCTTCTGGCTACCGCTGCTCTTCGCGCTCTCAGTGTCCTATTACACTGACCGTTGTC 1437
Qy      460  ProGluThrLysGlyThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db      1438  CCTGAGACTAAAGGCAGGACTCTGGAACAAGTTCACAGCCCATTTTCGAGGGACGA 1491
```

```
RESULT 2
AF289587
LOCUS      AF289587      1954 bp      mRNA      linear      HTC 01-JAN-2002
DEFINITION Homo sapiens clone pp7425 unknown mRNA.
ACCESSION  AF289587
VERSION     AF289587.1  GI:18027377
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE 1 (bases 1 to 1954)
AUTHORS   Zhang,P.P., Zhou,X.M., Jiang,H.Q., Huang,Y., Qin,W.X., Zhao,X.T.,
           Wan,D.F. and Gu,J.R.
TITLE      Novel human cDNA clones with function of inhibiting cancer cell
           growth
JOURNAL    Unpublished
REFERENCE 2 (bases 1 to 1954)
AUTHORS   Zhang,P.P., Zhou,X.M., Jiang,H.Q., Huang,Y., Qin,W.X., Zhao,X.T.,
           Wan,D.F. and Gu,J.R.
TITLE      Direct Submission
JOURNAL    Submitted (17-JUL-2000) National Laboratory For Oncogenes & Related
           Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Tu Road, Shanghai
           200032, P. R. China
FEATURES   Location/Qualifiers
            source
              1..1954
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="pp7425"
                /cdate="491.1237"
                /codon_start=1
                /product="unknown"
                /protein_id="AAL55771.1"
                /db_xref="GI:18027378"
                /translation="MVVVGILLAYLAWLWLVLCVPPSLMLLLMCMFETPR
                FLITQHRROEMALRELWGSEGWEDPPGAEQSFHLALRQPGLYKPTIGVSLNA
                FQQLGVNWFVTAETIFERAKFDSIASVVGVIQVLFVTAVALIMDRAGRLLIV
                LSGVWVFSTAFGAPFKLTQGGNGSSHVAISAPVSAQPDVADSLAWLAVGSMCLF
                IAGGPQALWSLLACFLHLQCPFFVLCP"
            CDS
              491..1237
                /codon_start=1
                /product="unknown"
                /protein_id="AAL55771.1"
                /db_xref="GI:18027378"
                /translation="MVVVGILLAYLAWLWLVLCVPPSLMLLLMCMFETPR
                FLITQHRROEMALRELWGSEGWEDPPGAEQSFHLALRQPGLYKPTIGVSLNA
                FQQLGVNWFVTAETIFERAKFDSIASVVGVIQVLFVTAVALIMDRAGRLLIV
                LSGVWVFSTAFGAPFKLTQGGNGSSHVAISAPVSAQPDVADSLAWLAVGSMCLF
                IAGGPQALWSLLACFLHLQCPFFVLCP"
ORIGIN
Alignment Scores: 5.29e-182 Length: 1954
Pred. NO.: 2123.50 Matches: 426
Percent Similarity: 89.14% Conservative: 1
Best Local Similarity: 88.94% Mismatches: 1
Query Match: 86.43% Indels: 51
DB: 11 Gaps: 1
US-09-886-954A-1 (1-477) x AF289587 (1-1954)
Qy      1  MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db      1  GTGAGCGCCGAGACCCAGAGAAACCCAGCGCTTCTGGGGCTCTCTGGCGGCGAGCGCG 60
Qy      21  ProArgGlyArgArgValPheLeuAlaPheAlaAlaLeuGlyProLeuSerPhe 40
Db      61  CCCCGCGCGCGCGCTTCTCGCGCGCTTCTCGCGCTTCTCGCGCGCTTCTCGCGCGCTTCTCG 120
Qy      41  GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
Db      121  GGCTTCGCGCTCGCTCAGCTCCCGCGCATCTCTAGCTGAGCGCGCGCGCGCGCGCG 180
Qy      61  AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
Db      181  GCGCGCGCGCTGGAGCGCGCGCGCTCTCTGGTTCGGGGCTGTCTGACCTGGGTGCC 240
Qy      81  AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db      241  CGCGCGGCGGAGTGTGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Qy      101  LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db      301  TTGTGTGCTCCGCGCGCTTCTCGTGGCGCTTCTCGCGCTTCTCGCGCTTCTCGCGCTTCTCG 360
Qy      121  TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db      361  TGGATGTCTGGGGGCGCGCTCTCTCACCGCGCTGGCTGGGTGTGTGTGTGTGTGTGTGTGTGT 420
Qy      141  AlaProValTyrIleSerGluIleAlaTyrProAlaVal-ArgGlyLeuLeuGlySerCys 160
Db      421  GCGCGCGTCTACATCTCCGAAATCGCTTACCCAGCAGTCCCGGGGTTGTCTCGGCTCTCTG 480
```

```

QY 160 sValGlnLeuMetValValValGlyLeuLeuAlaTyrLeuAlaGlyTrpValLeuG1 180
Db 481 TGTGCAGCTAATGGTCGTGGCATCTCTCTGGCTACCTGGCAGGCTGGTGTCTGA 540

QY 180 uTrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuLeuLeuMetCys 200
Db 541 GTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 600

QY 200 sPheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAl 220
Db 601 CTTCTATGCCAGACCCCGCGTCTCTGCTGACTCAGACAGCGCCAGAGGCATGGC 660

QY 220 aAlaLeuArgPheLeuTrpGlySerGlnGlyTrpGluAspProProIleGlyAlaG1 240
Db 661 CGCCCTGGGTTCTCTGGGCTCGAGCAGAGCTGGGAAGACCCCCCATCGGGCTGA 720

QY 240 uGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleG1 260
Db 721 GCAGAGCTTTCACCTGGGCCCTCTGGCGAGCCCGCATCTACAAGCCCTTCATCATGG 780

QY 260 yValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaG1 280
Db 781 CGTCTCCCTGATGGCTTCCAGCAGCTGTGGGGGTCAACGCCGTCATGTTCTATGCAGA 840

QY 280 uThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyVa 300
Db 841 GACCATCTTTGAAGAGGCCAAGTTCAAGACAGCAGCCTGGCTCGGTCTGTGGTGT 900

QY 300 lIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgG1e 320
Db 901 CATCCAGGTGCTGTTCACAGCTGTGGCGCTCTCATCATGGACAGCAGGCGGAGGCT 960

QY 320 uLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPh 340
Db 961 GCTCTGTGTCTGTTCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020

QY 340 eLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSe 360
Db 1021 CAAGCTGACCCAGGGTGGCCCTGGCAACTCTCTCGCACGTGGCCATCTCGGGCCCTCTC 1080

QY 360 rAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLe 380
Db 1081 TGCACAGCTGTGTATGCAGCGTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1140

QY 380 uPheIleAlaGly-PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluI 400
Db 1141 CTTCTATCGCGGAGTCTC----- 1157

QY 400 lePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuM 420
Db 1157 ----- 1157

QY 420 etAlaPheLeuValThrLysGluPheSerLeuMetGluValLeuArgProTyrGlyVa 440
Db 1158 -----CTCAGGCCCTATGGAG 1173

QY 440 laPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValP 460
Db 1174 CTTCTGTGCTTGGCTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1233

QY 460 roGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 1234 CTGAACCTAAAGGAAGACTCTGGAAACAAATCACAGCCCAITTTTGAGGGGCGA 1286

RESULT 3
AY414180 1008 bp DNA linear GSS 17-DEC-2003
LOCUS Homo sapiens SLC2A8 gene, VIRUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY414180
VERSION AY414180.1 GI:39770142
KEYWORDS GSS.
SOURCE Homo sapiens (human)

```

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 1008)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

Inferring nonneutral evolution from human-chimp-mouse orthologous

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

2 (bases 1 to 1008)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE

JOURNAL

COMMENT

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES

source

1..1008
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1008
/gene="SLC2A8"
/locus_tag="HQM5138"

gene

ORIGIN

Alignment Scores:

Pred. No.: 4,13e-145 Length: 1008
Score: 1713.00 Matches: 332
Percent Similarity: 99.10% Conservative: 0
Best Local Similarity: 99.10% Mismatches: 3
Query Match: 69.72% Indels: 0
DB: 29 Gaps: 0

US-09-886-954A-1 (1-477) x AY414180 (1-1008)

QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
Db 1 GTCTACATCTCGAATCGCTACCCAGCAGTCCGGGGTTCCTCGGTCTCTGTGTGAG 60
QY 163 LeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArg 182
Db 61 CTAATGGT 120
QY 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
Db 121 TGGGT 180
QY 203 ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaLeu 222
Db 181 CCCGAGACCCCGCGCTTCTCTGACTCAGACAGGCCAGGAGGCCATCGCGCCCTG 240
QY 223 ArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGluGlnSer 242
Db 241 CGGTTCCTGTGTGGGCTCCGAGCAGGGTGGGAAGACCCCCCATCGGGGTGTGAGCAG 300
QY 243 PheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGlyValSer 262
Db 301 TTTTACCTGGCCCTGTGTGGGACGCCCGGCATCTACAAGCCCTTCATCATCGGGGTCTCC 360
QY 263 LeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIle 282
Db 361 CTGATGGCTTCCAGCAGCTGTGGGGGTCAACCGCTCATGTTCTATGAGAGACCATC 420
QY 283 PheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyValIleGln 302
Db 421 TTTGAAGAGGCCAAGTTCAAGGACAGCAGCCTGGCTCGTGTGTGTGTGTGTGTGTGT 480

303 ValLeuPheThrAlaValAlaAlaLeuMetAspArgAlaGlyArgLeuLeu 322
 Db GTGCTGTTTCAGCTGGCGGCTCTCATCATGACAGAGAGGCGGAGGCTCTCTG 540
 323 ValLeuSerGlyValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLeu 342
 Db GTCTNNNNNGTGGTTCATGGTGTTCAGCACGAGTGCCITTCGGCGCTACTTCAAGCTG 600
 343 ThrGlnGlyProGlyAsnSerSerHisValAlaLeuSerAlaProValSerAlaGln 362
 Db ACCCAGGAGTGGCCCTGGCAACTCTCGCACGTGGCCATCTCGGCGCCTGTCTCTGCACAG 660
 363 ProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIle 382
 Db CCTGTTGATGCCAGCTGGGGCTGGCTGGCTGGCGTGGCAGCATGTGCTCTTCATC 720
 383 AlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePhePro 402
 Db GCCGCTTTGGGCTGGGCTGGGGGCCCATCCCTGGCTCCTCATGTCCAGATCTTCCT 780
 403 LeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPhe 422
 Db CTGCATGTCAGGCGGTGGCACAGCATCTCGCTCCTCACCAACTGGCTCATGGCCTTT 840
 423 LeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrp 442
 Db CTCGTGACCAAGGAGTTTCAGAGGCTCATGAGGTCCTCAGGCCCTATGGAGCCTTCCTGG 900
 443 LeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThr 462
 Db CTTCGCTCCGCTTTCTGCATCTTCAGTCTCCTTTTCACTTTGTTCTGTGCTCCCTGAACT 960
 463 LysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 Db AAGGAAAGACTCTGGAACAAATCACAGCCCAATTTGAGGGGCGA 1005

RESULT 4

LOCUS BX395379 1201 bp mRNA linear EST 13-MAY-2003
 DEFINITION BX395379 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
 cDNA clone CS0DD006YP06 5-PRIME, mRNA sequence.
 ACCESSION BX395379
 VERSION BX395379.1 GI:30620615
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT Genoscope
 COMMENT Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5383.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DD006DH03QP1&cluster=5383.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DD006DH03QP1.
 Location/Qualifiers
 FEATURES
 source
 1. 1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DD006YP06"
 /tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:
 Pred. No.: 9, 65e-125 Length: 1201
 Score: 1490.50 Matches: 320
 Percent Similarity: 68.72% Conservative: 3
 Best Local Similarity: 68.09% Mismatches: 23
 Query Match: 60.66% Indels: 127
 DB: 13 Gaps: 1
 US-09-886-954A-1 (1-477) x BX395379 (1-1201)
 Qy 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 Db ATGACGCCCGAGGACCCAGAGAAACCCAGCGCTTCTGGGGCTCTCTGGCGGCGAG--- 138
 Qy 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaLeuGlyProLeuSerPhe 40
 Db 138 ----- 138
 Qy 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 Db 138 ----- 138
 Qy 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 Db 138 ----- 138
 Qy 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 Db 138 ----- 138
 Qy 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 Db 138 ----- 138
 Qy 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 Db 138 ----- 138
 Qy 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 Db 139 -----GTCTACATCTCCGAAATCGCTACCCAGCAGTTCGCGGGGTTGCTCGGCTCCTGT 192
 Qy 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 Db 193 GTGCAGCTAATGTCTGCTCGGCATCTATG-GCCTACCTGGCAGGCTGGTGTGTTGGAG 251
 Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 Db 252 TKGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 311
 Qy 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgGlnGlnAlaMetAla 220
 Db 312 TTTATGCCGAGAGCCCGCGCTCTCTGACTCAGCACAGGCGCCAGGAGGCGCATGGCC 371
 Qy 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
 Db 372 GCCTTGGGNTTCTGTGGGCTCCGAGCAGGCTGGGAAGACCCCTCCATCGGGGCTGAG 431
 Qy 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGly 260
 Db 432 CAGAGCTTTCACCTGGCCCTGCTGGCGCAGCNCGGCATCTACAGCYCTTHATCATCGGC 491
 Qy 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
 Db 492 GTCTCCCTGATGGCTTCCAGCAGCTGTGGGGGTTAAACGCGGTATGTTCTATGACAGAG 551
 Qy 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300

```

Db 552 ACCATCTTTGAAGAGGCCCAAGTTCAAGGACAGACGACGCTGGCTCGTCTGGTGGTGT 611
Qy 301 lleGlnValLeuPheThAlaValAlaAlaLeuIleMetAspArgAlaGlyArgLeu 320
Db 612 ATCCAGGTGCTGTWCACAGCTGTATCGGCTCTYATATGGACAGACGAGCGGCGAGGCTG 671
Qy 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
Db 672 CTCCTGGTCTGTGTYAGGTGTGTGTCATGTGTTCACACAGAGTTCCTTGGCGCCCTACTTC 731
Qy 341 LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
Db 732 AAGCTGACCCAGGGTGGCCCTGGCAACTCCTCGCACGTGWCACATWTMGCGGCTCTTCTT 791
Qy 361 AlaGlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValGlySerMetCysLeu 380
Db 792 GCACAGCTGTGTGATGCCAGCTGGGGCTGGCTGGCTGGCGGTGGGACGATGTGSCCTW 851
Qy 381 PheIleAlaGlyPheAlaValGlyTyrGlyProIleProTyrLeuLeuMetSerGluIle 400
Db 852 TWATCGCGCGCTTTGGGTGGCTGGGGCCCATCCCTGGCTCTCATGTAGAAATC 911
Qy 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
Db 912 TTCCTCTCGCATGTTAAGGGCGTGGCGACAGGATCTGGCTTCTCACCACTGGGCTCATG 971
Qy 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
Db 972 GCCTTTCTCGTGACCAAGAGTTCACAGCTTCATGAGGCTCTCAGGCCCTATGGAGCC 1031
Qy 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
Db 1032 TTCYKGTCTKCTCCGCTTTCTGCACTCCAGTTCCTTTTCACCTTTGTTCTKTTGTCT 1090
Qy 461 -GluThrLysGlyLysThrLeuGluGln 469
Db 1091 TAAAACTAAAGAAATTTTGGACAA 1118

RESULT 5
LOCUS AY414182 1000 bp DNA linear GSS 17-DEC-2003
DEFINITION Mus musculus SLC2A8 gene, VIRTUAL TRANSCRIPT, partial sequence,
ACCESSION AY414182
VERSION AY414182.1 GI:39770144
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Gargill,M.
Infering nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Gargill,M.
Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..1000
/organism="Mus musculus"

```

```

/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1000
/gene="SLC2A8"
/locus_tag="HCW5138"

ORIGIN
Alignment Scores: 2.05e-124 Length: 1000
Pred. No.: 1485.50 Matches: 284
Score: 91.07% Conservative: 22
Percent Similarity: 84.52% Mismatches: 26
Best Local Similarity: 60.46% Indels: 4
Query Match: 29 Gaps: 2
DB:

US-09-886-954A-1 (1-477) x AY414182 (1-1000)
Qy 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
Db 1 GTTTACATCTCGAAATCGCTTACCCAGCTGCCGAGACTGCTCGGCTCTCTGTGTGCGAG 60
Qy 163 LeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGluTyrparg 182
Db 61 CTGATGGTGTGTCACCTGGCATCTCTCTGGCCTATGTGGCAGGCTGGGTCTTAGAGTGGGC 120
Qy 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
Db 121 TGGCTGGCGGTGTGGGCTGTGGCCCCCACCCTCATGCTGTGCTCATGTGCTACATG 180
Qy 203 ProGluThrProArgPheLeuLeuThrGlnHisArgGlnGluAlaMetAlaAlaLeu 222
Db 181 CCCGAGACCCACAGTTCCTCTCCTCACCAACACACGACGAGGAGGCCATGGTGCCTTG 240
Qy 223 ArgPheLeuTrpGlySerGluGlnGlyTyrGluAspProProIleGlyAlaGluGlnSer 242
Db 241 CGCTTCCTGTGGGCTCTGAGGAGGCGTGGGAAGAGCCCTGTTGGGGCTGAGCACGCGC 300
Qy 243 PheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGlyValSer 262
Db 301 TTCAGCTGGCCCTGTCTGAGGCGCCCTGGGCACTACAGCCCTCATCATCGGCATTTCC 360
Qy 263 LeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIle 282
Db 361 CTCATGGTCTTCCAGCAGCTGTGAGGGTCAATGCTATCATGTTCTATGCCAACAGCATC 420
Qy 283 PheGluGluAlaLysPheLysAspSerSer-LeuAlaSerValValValGlyValIleG1 302
Db 421 TTCAGGAGGCCAAAGTTCAAG-----CAGCCTGGCCTCGGTCACTGTGGGCATAATCCA 474
Qy 302 nValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuLe 322
Db 475 GGTCTGTTCATCTGCTGGCGGCCCTCATCATGACAGACGAGGCGGAGGCTGCTCT 534
Qy 322 uValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLe 342
Db 535 GGCCTTGTGCGGTGTGATCATGGTGTTCAGTATGAGTGCCTTTGGTACTACTTCAAACT 594
Qy 342 uThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAlaG1 362
Db 595 GACCCAGAGCCTCCCCAGCAACTCTCTCCACGTAGGCTG---GTGCCCATCTCGCGCGGA 651
Qy 362 nProValAspAlaSerValGlyLeuAlaTyrLeuAlaValGlySerMetCysLeuPheI1 382
Db 652 GCCTGTGATGTCCAAGTGGACATGGCTGGCTGGCTGGCTGGTAGCAGCATGTGCTCTTCAT 711
Qy 382 eAlaGlyPheAlaValGlyTyrGlyProIleProTyrLeuLeuMetSerGluIlePhePr 402
Db 712 TGCTGGCTTTTGGGTGGGCTGGGACCATCCCTGGCTCCTCATGTGACAGATCTTCCC 771
Qy 402 oLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPh 422
Db 772 TCTGCATGTCAAGGGTGTGGCTACCGGCATCTGTGTCTCTCACCACACTGTTTCATGGCTT 831

```

```

Qy 422 eLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTr 442
Db 832 TCTAGTGACCAAGAGATTACAGCGTCATGGAGATGCTCAGACCTACGGTGGCTTCTG 891
Qy 442 pLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluTh 462
Db 892 GCTCACCCTGCTTCTGCGCTCTCAGTGTCTCTATTACACTGACCGGTGTCCTCGTAGAC 951
Qy 462 rLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 952 TAAAGGCAGGACTCTGGAAACAAGTCACAGCCCAATTTCGAGGGACGA 997

RESULT 6
BM545247
LOCUS BM545247 1085 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_6497266 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5726945
5', mRNA sequence.
ACCESSION BM545247
VERSION BM545247.1 GI:18777177
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1085)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12720 row: d column: 18
High quality sequence stop: 679.
Location/Qualifiers
1. 1085
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5726945"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_124"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcorV
(destroyed); Site 2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcorV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."

ORIGIN
Alignment Scores:
Pred. No.: 1.15e-114 Length: 1085
Score: 1379.00 Matches: 279
Percent Similarity: 87.77% Conservative: 1
Best Local Similarity: 87.46% Mismatches: 6
Query Match: 56.13% Indels: 33
DB: 12 Gaps: 3
US-09-886-954A-1 (1-477) x BM545247 (1-1085)

Qy 162 GlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGluTyrP 181
Db 1 CAGCTAAATGGTGGTGGGATCCTCTGGCTACCTGGCGAGGCTGGGTGGTGGAGTGG 60
Qy 182 ArgTTrpLeuAlaValLeuGlyCysValProSerLeuMetLeuLeuMetCysPhe 201

```

```

Db 61 CGCTGGCTGGCTGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 120
Qy 202 MetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAlaAla 221
Db 121 ATGCCCGAGAGCCCGCGCTTCTGCTGACTCAGCACAGGCGCCAGAGGCGCATGGCCGCC 180
Qy 222 LeuArgPheLeuTyrGlySerGluGlnGlyTyrGluAspProProIleGlyValAlaGlu 241
Db 181 CTGGGGTTCCTGTGGGGCTCCGAGCAGGGCTGGGAAGACCCCTCATCGGGGCTGAGCAG 240
Qy 242 SerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGlyVal 261
Db 241 AGCTTTTCACTGGCCCTGCTGGCGCAGCCCGCATCTACAAGCCCTTCATCATCGGGCTC 300
Qy 262 SerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThr 281
Db 301 TCCCTGATGGCTTCCAGCAGCTGTCGGGGTCAACGCGTTCATGTTCTATGCAGAGACC 360
Qy 282 IlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyValIle 301
Db 361 ATCTTTTGAAGAGGCCAAGTTCAAGCAGCAGCCTGGCTCGTGGTGGTGGTGGTGGTGGT 420
Qy 302 GlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeu 321
Db 421 CAGTGTCTGTTCACAGCTGTGGCGCTCTCATCATGACAGCAGCGCGGAGGCTGCTC 480
Qy 322 LeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLys 341
Db 481 CTGGTCTTGTTCAGGTGGTTCATGGTTCAGCAGCAGTGGCTTCGGCGGCTTCTCAAG 540
Qy 342 LeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAla 361
Db 541 CTGACCCAGGCTGGCCCTTGGCACTCTTCGCACGTGGCCATCTCGGCGCTTCTCTGCA 600
Qy 362 GlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValGlySerMetCysLeuPhe 381
Db 601 CAGCTGTGTGATGCCAGGTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 660
Qy 382 IleAlaGlyPheAlaVal-GlyTyrGlyProIleProIleProIleLeuMetSerGluIlePh 401
Db 661 ATCGCCGGCTTTGGGTTGGCTGGGCGCCATCCCTGGCTCTCATGTTCAGAGATCTT 720
Qy 401 eProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrLeuMetAl 421
Db 721 CCCTCTCATGTCAAGGGCGTGGCGACAGGCATCTGGTCTCTCCACCACTGGCTCATGGC 780
Qy 421 aPheLeuValThrLysGluPheSerSerLeuMet----- 432
Db 781 CTTTCTCTGACCAAGAGATTACAGACCTCATGCTCTGCTGGCTCTGAGGACTCAN 840
Qy 433 -----GluValLeuArgPro----- 437
Db 841 GAACACCTTCAGCTTTGCAAGACCTCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 900
Qy 438 -----TyrGly-----AlaPheTyrLeuAlaSerAlaPhe 447
Db 901 NAAAAAGAGGGGGGCGCTCCAAAGGACCTTTGGCTTCTGGCTTGAAGGGCGCTTTT 955

RESULT 7
BM5479842
LOCUS BM5479842 932 bp mRNA linear EST 21-MAR-2001
DEFINITION 602527373F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4650906 5',
mRNA sequence.
ACCESSION BM5479842
VERSION BM5479842.1 GI:13412121
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 932)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

```

TITLE
JOURNAL
COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW1433 row: a column: 19
High quality sequence stop: 820.

FEATURES

source

1. 932
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4650906"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_21"
/note="Organ: placenta; Vector: pORF7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:

Pred. No.: 3 18e-111 Length: 932
Score: 1340.00 Matches: 293
Percent Similarity: 93.93% Conservative: 1
Best Local Similarity: 93.61% Mismatches: 17
Query Match: 54.54% Indels: 7
DB: 12 Gaps: 0

US-09-886-954A-1 (1-477) x BG479842 (1-932)

QY 51 IleProSerLeuGlnArgAlaAlaProProAlaProArgLeuAspAlaAlaSer 70
DB 2 ATCCCTAGCCTGCAGCGCGCGCGCCCGCGCGCGCGCGCGCGCGCGCTCC 61
QY 71 TrpPheGlyAlaValThrLeuGlyAlaAlaGlyGlyValLeuGlyGlyTrpLeu 90
DB 62 TGGTTTCGGGGTGTCTGACCCCTGGGTCGCGCGCGCGCGCGCGCGCTGG 121
QY 91 ValAspArgAlaGlyArgLysLeuSerLeuLeuLeuCysSerValProPheValAlaGly 110
DB 122 GTGACCG 181
QY 111 PheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGlyGlyArgLeuThr 130
DB 182 TTTGCGGTCATCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 241
QY 131 GlyLeuAlaCysGlyValAlaSerLeuValAlaProValTrpIleSerGluIleAlaTrp 150
DB 242 GGCCTGGCGCTCGGGGTGTGCTCCCTAGTGGCGCGCGCGCGCGCGCGCGCTAC 301
QY 151 ProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetValValGlyIleLeu 170
DB 302 CCAGCAGTCCGGGGGTGTCTGGCTTCCTGTGTGACGCTAATGCTGTGGCATCTCT 361
QY 171 LeuAlaTrpLeuAlaGlyTrpValLeuGluTrpArgTrpLeuAlaValLeuGlyCysVal 190
DB 362 CTGGCTACCTCGGCGAGGTGGTGTGGAGTGGCGCTGGCTGGCTGGCTGGCTGGCT 421
QY 191 ProProSerLeuMetLeuLeuLeuMetCysPheMetProGluThrProArgPheLeuLeu 210
DB 422 CCCCCCTCCCTCATGCTGCTCTCATATGCTTCTATGCTTCTATGCTGCTGCTGCTG 481

QY 211 ThrGlnHisArgArgGlnGlnAlaMetAlaAlaLeuArgPheLeuTrpGlySer-GluG 230
DB 482 ACTCAGCACAGCGCCAGAGGCCATGGCCGCCCTTCCCTGTGGGGTCTCCTGAGCA 541
QY 230 nGlyTrpGluAspProProIleGlyValAlaGlnSerPheHisLeuAlaLeuLeuArgG 250
DB 542 GGGCTGGGAGAGCCCGCCCATCGGGGCTGAGCAGAGCTTTCACCTGGCCCTGCTGGGCA 601
QY 250 nProGlyIleTrpLysProPheIleGlyValSerLeuMetAlaPheGlnGlnLeuSe 270
DB 602 GCGCGCATCTACAAGCCCTTCATCATCGGGCTCTCCCTCATGGGCTTCCAGCAGCTGTC 661
QY 270 rGlyValAsnAlaValMetPheTyrAlaGluThrIlePheGluGluAlaLysPheLysAs 290
DB 662 GGGGTCAACGCCGTCATGTTCTATGACAGACCATCTTTGAAGAGGCCAGATTCAAGA 721
QY 290 pSerSerLeuAla-SerValValValGlyValIleGlnValLeuPheThrAlaValAlaA 310
DB 722 CAGCAGCCTGGCCCTCGCTCGTGTGCTGTCATCCAGGTGCTGTTTCACAGCTGTGGGG 781
QY 310 laLeuIleMetAspArgAlaGlyArgLeuLeuLeuValLeuSerGlyValValMetV 330
DB 782 CTCTCATCATGGACAGACAGCGCGG-AGGCTGGTCTCTGCTGG-TCAGGTGTGGCATGG 839
QY 330 alPheSerThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyGlyProGlyAsnS 350
DB 840 TGTACGGCGGAGTGCTTC-GCGGCTTAATTCAAGCTGACCCG-GGTTGGCGCTGCAATT 897
QY 350 erSerHisValAlaIleSerAlaProValSerAla 361
DB 898 CCTCGCC-GTGGCCATCTCGGGCCTTGTTCCTGG 931

RESULT 8

BG700749

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

830 bp mRNA linear EST
602681616F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4814451 5',
mRNA sequence.
BG700749 GI:13970402
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 830)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW10709 row: p column: 04
High quality sequence stop: 789.

FEATURES

source

1. 830
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4814451"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_95"
/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer


```
Db 303 CTGGCCCTGCTGGCGCAGCCCGGCATCTACAAGCCCTTCGTATCGGCGCTCTCCCTGATG 362
Qy 265 AlaPheGlnGlnSerGlyValAenAlaValMetPheTyrAlaGluThrIlePheGlu 284
Db 363 GCCTTCACAGCAGTGTGCGGGGTACACGGCGTCATGTTCTATGCAGAGACCATCTTTGAA 422
Qy 285 GluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyValIleGlnValLeu 304
Db 423 GAGGCCAGTTTCAGACAGCAGCGCTGGCTCGTATCGTGGTGTATCAGAGTGCTG 482
Qy 305 PheThrAlaValAlaLeuLeuMetAspArgAlaGlyArgLeuLeuValLeu 324
Db 483 TTCACAGCTGTGGCGCTCTCATCATGACAGACAGCGGAGGCTGTCTCTGTGTCNNG 542
Qy 325 SerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLeuThrGln 344
Db 543 TCAGGTGTGTATGTTGTTCAGACAGAGTGCCTTCGGCGCCTACTTCAAGCTGACCCAG 602
Qy 345 GlyGlyProGlyVasnsSerHisValAlaIleSerAlaProValSerAlaGlnProVal 364
Db 603 GTGGCCCTGGCAACTCTTCGCACAGTGGCCCTCTCGCGCCTGTCTCTGCACAGCCTGT 662
Qy 365 AspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAlaGly 384
Db 663 GATGCCAGCGTGGGGCTGGCTGGCTGGCGTGGCGGAGCATGTGCCTCTTCATCGCGGC 722
Qy 385 PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePheProLeuHis 404
Db 723 TTTGGCGTGGCTGGGGGCCCATCCCTGGCTCTCATGTACAGAGATCTTCCCTCTGTCAT 782
Qy 405 ValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeuVal 424
Db 783 GTCAAGGGCGTGGCAGAGCATCTCGCTCCACCACTGGCTATGGCCCTTTCGCG 842
Qy 425 ThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeuAla 444
Db 843 ACCAAGGAGTTCAGCAGCCTCATGAGAGTCTCTCAGGCCCTATGGAGCCTTCTGGCTGCC 902
Qy 445 SerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrLysGly 464
Db 903 TCCGCTTCTGTCATCTTCAGTGTCTCTTTTCACTTTGTGTCTCCCTGAAACTAAAGGA 962
Qy 465 LysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 963 AAGACTCTGGAACAAATCACGCCCAATTTTGAGGGCGGA 1001
```

```
RESULT 11
LOCUS CD557801
DEFINITION AGENCOURT_14413291_NIH_MGC_180_Homo_sapiens_CDNA_clone
IMAGE:30386912_5', mRNA sequence.
ACCESSION CD557801
VERSION CD557801.1 GI:31583869
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 875)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
```

FEATURES
source

```
http://image.llnl.gov
Place: ND4W456 row: f column: 09
High quality sequence stop: 618.
Location/Qualifiers
1..875
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30386912"
/lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
/clone_lib="NIH_MGC_180"
/note="Organ: Testis; Vector: pCMV-SPORT6.1; Site 1: NotI;
Site 2: EcoRV (destroyed); Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.68 kb. Library was
constructed by (Invitrogen). Note: this is a NIH_MGC
Library."
```

ORIGIN

```
Alignment Scores:
Pred. No.: 1..26e-104 Length: 875
Score: 1267.00 Matches: 265
Percent Similarity: 95.05% Conservative: 4
Best Local Similarity: 93.64% Mismatches: 7
Query Match: 51.57% Indels: 7
DB: 14 Gaps: 0

US-09-886-954A-1 (1-477) x CD557801 (1-875)

Qy 76 ValThrLeuGlyValAlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGly 95
Db 1 GTGACCTCGGGTGGCGCGGGAGTGTCTGGCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 60
Qy 96 ArgLysLeuSerLeuLeuLeuCysSerValProPheValAlaGlyPheAlaValIleThr 115
Db 61 CGCAAGCTGAGCCTCTGTGTGCTCGTCCGTGCGCTTCGTGGCGGCTTGTCCGCTCATCACC 120
Qy 116 AlaAlaGlnAspValTrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGly 135
Db 121 GCGGCCAGGAGCGTGGATGCTGTGGGGGGCCCTCTCACCAGGCTGGCTGGCTGGCTGG 180
Qy 136 ValAlaSerLeuValAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGly 155
Db 181 GTTGCTCTCCCTAGTGGCCCGGCTACATCTCCGAATCGCCTACCCAGAGTCCGCGGGG 240
Qy 156 LeuLeuGlySerCysValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAla 175
Db 241 TTGCTCGGCTCTCTGTGTGACGCTAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Qy 176 GlyTrpValLeuGluTrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMet 195
Db 301 GGTGGGTGTGTGGAGTGGCGTGGTGGTGTGTGTGGGTGTGTGTGGGTGTGTGTGGGTGTGT 360
Qy 196 LeuLeuLeuMetCysPheMetProGluThrProArgPheLeuLeuThrGlnHisArgArg 215
Db 361 CTGCTTCTCATGTGCTTATGCCCGAGACCCCGCGCTCTCTGTGTGTGTGTGTGTGTGTGTGT 420
Qy 216 GlnGluAlaMetAlaAlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspPro 235
Db 421 CAGGAGGCCATGGCCGCTCGGTTCCTGTGGGGTCCGAGCAGGGGTGGGAAGACCC 480
Qy 236 ProIleGlyAlaGluGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLys 255
Db 481 CCCATCGGGGCTGAGCAGAGCTTTTCACTGGCCCTGTGTGTGGTGTGTGTGTGTGTGTGTGT 540
Qy 256 ProPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaVal 275
Db 541 CCTTCATCATCGGGCTCTCCCTGATGCCCTTCCAGCAGCTGTCCGGGGTCAACCCGCTC 600
Qy 276 MetPheTyrAlaGluThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSer 295
Db 601 ATGTTCTATGAGAGACCATCTTTGAAGAGGCCAAGTTCAAGGACAGCAGGCTGGCCCTCG 660
```

```

QY 296 ValValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArg 315
Db 661 GTTCGTGGTGTCTATCCAGGTGCTGTTTCACAGCTGTGGCGCTCTCTCATGGACAAA 720
QY 316 AlaGlyArgArgLeuLeuValLeuSerGlyVal-ValMet-ValPheSerThrSer 334
Db 721 CCAGGCGGAGGCTGCTCTGGGTCTGTGTACAGGGTGGGTCTATGGGTGTTCAGCACAAAT 780
QY 335 Ala--PheGlyAlaTyrPheLysLeu--ThrGlnGlyGlyProGlyAsnSerSer 351
Db 781 GCCCTTTCGCCCTTACTTCAAGCTGAACCCAGGGTGGGCCCCCGGGCAAATTC 835

RESULT 12
LOCUS BM044230
DEFINITION 880 bp mRNA linear EST 07-NOV-2001
603621486F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5447181 5',
mRNA sequence.
ACCESSION BM044230
VERSION BM044230.1 GI:16773497
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://imgc.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1930 row: k column: 22
High quality sequence stop: 869.
FEATURES
    source
        1..880
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5447181"
            /tissue_type="carcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_40"
            /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
            Site 2: EcoRI; cDNA made by oligo-dT priming.
            Directionally cloned into EcoRI/XhoI sites using the
            following 5' adaptor: GGCAAGAG(G). Library constructed by
            Ling Hong in the laboratory of Gerald M. Rubin (University
            of California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies).
            Note: this is a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.: 2.67e-102 Length: 880
Score: 1241.50 Matches: 280
Percent Similarity: 58.82% Conservative: 0
Best Local Similarity: 58.82% Mismatches: 5
Query Match: 50.53% Indels: 193
DB: 12 Gaps: 2

US-09-886-954A-1 (1-477) x BM044230 (1-880)
QY 4 GluAspProGluGluThrGlnProLeuLeuGlyProGlySerAlaProArgGly 23
Db 4 GAGGACCCAGGGAACCCAGCGCTTCGGGGCTCTCTGGCGCAG-----50
QY 24 ArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAla 43

```

```

Db 50 -----50
QY 44 LeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProAlaProArg 63
Db 50 -----50
QY 64 LeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaGly 83
Db 50 -----50
QY 84 GlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuLeuCys 103
Db 50 -----50
QY 104 SerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeu 123
Db 50 -----50
QY 124 LeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaProVal 143
Db 51 -----GTC 53
QY 144 TyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGlnLeu 163
Db 54 TACATCTCCGAAATCGCCTACCCAGCAGTCCGGGGGTTCCTGGCTCCTGTGTGCAGCTA 113
QY 164 MetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArgTrp 183
Db 114 ATGGTCGTCTGGCATCTCTCTGGCTACCTGGCA-----149
QY 184 LeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMetPro 203
Db 149 -----149
QY 204 GluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaLeuArg 223
Db 149 -----149
QY 224 PheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGluGlnSer-Ph 243
Db 150 -----GAGCTT 155
QY 243 eHisLeuAlaLeuLeuArgGlnProGlyIleTyrIysProPheIleIleGlyValSerLe 263
Db 156 TCACCTGGCCCTGTCTGGCGCAGCCGGCATCTACAAGCCCTTCATCATCGCGCTCCCT 215
QY 263 uMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIlePh 283
Db 216 GATGGCTTCCAGCAGCTGTGGGGGTCAACGCCGTCATGTTCTATGCAGAGACCATCTT 275
QY 283 eGluGluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyValIleGlnVa 303
Db 276 TGAAGAGGCCAAGTTCCAGACAGCAGCTGGCTCGCTCGTGGGTGTCTATCAGGT 335
QY 303 lLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgLeuLeuLeuVa 323
Db 336 GCTGTTTCACAGCTGTGGCGCTCTCATCATGACAGCAGCGCGGAGGCTGCTCTGCTGT 395
QY 323 lLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLeuTh 343
Db 396 CTTGTGAGGTGTGTGTCATGTTTCAGCAGCAGTGTTCGGCGCCTACTTCAAGCTGAC 455
QY 343 rGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAlaGlnPr 363
Db 456 CCAGGGTGGCCCTGGCAACTCTCTCGACAGTGGCCATCTGGCGCCTGCTCTGCACAGCC 515
QY 363 oValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAl 383
Db 516 TGTGTGATCCAGCGTGGGGCTGGCTGGCTGGCGGTGGCGCAGCATGTGCTCTTCATCGC 575
QY 383 aGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePheProLe 403

```

Db 576 CCGCTTGGCGTGGCGTGGGGGCCATCCCTCGGCTCCTCATGTACAGAGATCTTCCTCT 633
 Qy 403 uHisValIysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLe 423
 Db 636 GCATGTCAAGGGCGTGGCGACAGCACTGGGTCTCTACCAACTGGCTCATGGCGCTTCT 695
 Qy 423 uValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLe 443
 Db 696 CGTGACCAAGAGGTTTCAGACGCCTCATGGAGGTCCTCAGGCGCTTATGGAGCGCTTCTGGCT 755
 Qy 443 uAlaSerAlaPheCysIlePheSerValLeuPheThrIleuPheCys-ValProGluThrL 463
 Db 756 TGGCTCCGTTTCTGCATTCCTCAGTGCCCTTTTCACTTTGTCTGGTGTCCCTGAAACTA 815
 Qy 463 YSGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 Db 816 AAGGAAGACTCTGAACAATCACGCCA--TTTGAGGGCGCA 857

 RESULT 13
 BG478000
 LOCUS
 DEFINITION 602522606F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4641145 5', mRNA linear EST 21-MAR-2001
 mRNA sequence.
 ACCESSION BG478000
 VERSION BG478000.1 GI:13410279
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 828)
 NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCMI407 row: k column: 02
 High quality sequence stop: 788.
 Location/Qualifiers
 1..828
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4641145"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_20"
 /note="organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.04e-102 Length: 828
 Score: 1238.00 Matches: 250
 Percent Similarity: 95.08% Conservative: 1
 Best Local Similarity: 94.70% Mismatches: 8
 Query Match: 50.39% Indels: 5
 DB: 12 Gaps: 0
 US-09-886-954A-1 (1-477) x BG478000 (1-828)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10728 row: i column: 23
 High quality sequence stop: 722.
 Location/Qualifiers
 1. .777
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4821598"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 97"
 /note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dt primed using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 1.3e-101 Length: 777
 Score: 1233.00 Matches: 243
 Percent Similarity: 98.39% Conservative: 1
 Best Local Similarity: 97.98% Mismatches: 3
 Query Match: 50.18% Indels: 2
 DB: 12 Gaps: 0

US-09-886-954A-1 (1-477) x BG171034 (1-777)

QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 DB 36 ATGACGCCCGGAGACCCAGAGAAACCCAGCGCTTCTGGGGCTCTCTGGCGCAGCGG 95
 QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 DB 96 CCCGCGCGCGCGCGCTTCTCTCGCGCTTCTCGCGCTCGCGCTCGCGCTCGCGCTTC 155
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 DB 156 GGGTTTGGCGTGGCTACAGTCCCGCGCATCCCTAGCTGAGCGCGCGCGCGCGCGCG 215
 QY 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 DB 216 GCCCGCGCGCTGACGACGCGCGCGCTCTCTGGTGGCGCTGCTGGTGGCGTGGC 275
 QY 81 AlaAlaGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 DB 276 GCGCGCGGGGAGTGTGGCGCGCTGGCTGGTGGACGCGCGCGCGCGCGAGCTGAGCCTC 335
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 DB 336 TTGCTGTCTCGTGGCGCTTCTGGCGCGCTTCTGGCGCTATCACCGCGCGCGCGAGCGTG 395
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 DB 396 TGGATGCTGTGGGGGCGCGCTCTCACCGCGCTGGCGCTGTGGCTCCCTAGTG 455
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 DB 456 GCCCGCGGTCTACATCTCCGAAATGGCTACCCAGCAGTCCGGGGGTGTCTCGGCTCTGT 515

QY 161 ValGlnLeuMervValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 DB 516 GTGCAGCTAATGTCGTCTGCGCATCTCTGGCTACTGCGAGGCTGGGTGCTGGAG 575
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 DB 576 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 635
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 DB 636 TTCATGCCGAGACCCCGGCTTCTGCTGACTCAGCACAGGCGCGAGGCGCATGGCC 695
 QY 221 AlaLeuArgPheLeu-TripGlySerGluGlnGlyTrpGluAspProProLeuGlyAlaG 240
 DB 696 GCCCTGCGGTTCCTGTGGGGTCCGAGCATGGTGGGAGAGACCCCATC-GGAGCTGA 754
 QY 240 uGlnSerPheHisLeuAlaLeu 247
 DB 755 CCAGAGCTTCCACCTGGACCTG 776

RESULT 15

BI757409

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1. .889

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5199778"

/lab_host="DH10B"

/clone_lib="NIH MGC 114"

/note="Organ: brain; Vector: pcMV-SPORT6; Site 1: NotI;

Site 2: EcoRV (destroyed); RNA source anonymous pool of 6

male brains, age range 23-27 yo. Library is oligo-dt

primed and directionally cloned (EcoRV site is destroyed

upon cloning). Average insert size 1.5 kb, insert size

range 1-3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics tracking code 019. Note:

this is a NIH_MGC Library."

FEATURES

source

Location/Qualifiers

1. .889

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5199778"

/lab_host="DH10B"

/clone_lib="NIH MGC 114"

/note="Organ: brain; Vector: pcMV-SPORT6; Site 1: NotI;

Site 2: EcoRV (destroyed); RNA source anonymous pool of 6

male brains, age range 23-27 yo. Library is oligo-dt

primed and directionally cloned (EcoRV site is destroyed

upon cloning). Average insert size 1.5 kb, insert size

range 1-3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics tracking code 019. Note:

this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 2.72e-101 Length: 889
 Score: 1230.50 Matches: 265
 Percent Similarity: 78.87% Conservative: 0
 Best Local Similarity: 78.87% Mismatches: 4
 Query Match: 50.08% Indels: 69

Search completed: September 28, 2004, 04:16:06
Job time : 2866 secs

```
DB: 12 Gaps: 1
US-09-886-954A-1 (1-477) x BI757409 (1-889)
QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
Db 38 GTCTACATCTCGAAATCGCTACCCAGCAGTCGGGGGTTGCTCGGCTCCGTGTGCAG 97
QY 163 LeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArg 182
Db 98 CTAATGGTCGTGCGGCATCTCTGGCCTACCTGGCA----- 136
QY 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
Db 136 ----- 136
QY 203 ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaLeu 222
Db 136 ----- 136
QY 223 ArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGluGlnSer 242
Db 137 -----GAG 139
QY 243 -PheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGlyValSe 262
Db 140 CTTTACCTGGCCCTGCTGCGGAGCCCGGCATCTACAAGCCCTTCATCATCGGGTCTC 199
QY 262 rLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIl 282
Db 200 CCGTATGGCCTTCCAGCAGCTGTCGGGGGTCAAGCCGTCATGTTCTATGCAGAGACCAT 259
QY 282 ePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyValIleGl 302
Db 260 CTTTGAAGAGGCAAGTTCAAGACAGACAGCCCTGGCCTCGTCGTCGTGGTGTATCCA 319
QY 302 nValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeu 322
Db 320 GGTGCTGTTCACAGCTGTGGCGCTCTCATATGAGACAGAGGCGGAGGCTGCTCCT 379
QY 322 uValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLe 342
Db 380 GGTCTGTGTGAGGTGTGTCATGGTGTTCAGCAGAGTGCTTCGGGCGCTACTTCAAGCT 439
QY 342 uThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAlaGl 362
Db 440 GACCCA-GGTGGCCCTGGCAACTCCTCGCACGTGGCCATCTCGGCGCCTGTCTCTGCACA 498
QY 362 nProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIl 382
Db 499 GCCTGTTGATGTCAGCGTGGGGCTGGCCTGGCTGGCCGTTGGCGAGCATGTGCTCTTCAT 558
QY 382 eAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePhePr 402
Db 559 CGCGGGCTTTGGGTGGGTGGGGGCCCATCCCTGGCTCCTCATGTCAAGATCTTCCC 618
QY 402 oLeuHisValLysGlyValAlaAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPh 422
Db 619 TCTGCATGTCAAGGGGCTGGCGCAGGCATCTGCGTCCTCACCAACTGGCTCATGGCCTT 678
QY 422 eLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTr 442
Db 679 TCTCGTGACCAAGAGGCTTCAGCAGCCCTCATGGAGGTCCTCAGGCCCTATGGAGCCTTCG 738
QY 442 pLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluTh 462
Db 739 GCTTGCTCCG--TTTCTGATCTTCAGTGTCTTTTCACTTTGTTCTGTGTCCTGAAC 796
QY 462 rLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 797 TAAAGGAAAGACTCTCGAAACAATCACAGCCCATTTTGGGGCGCA 842
```

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: September 28, 2004, 00:52:18 ; Search time 484 Seconds
(without alignments)
4186.757 Million cell updates/sec
Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPEDPRTQPLGGPGGSA.....CVPETKTKLTQITAHFEGR 477

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO spool_p/US09886954/runat_27092004_170631_13378/app_query.fasta_1.647
-DB=N_Geneseq_29Jan04 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09886954.ecgn_1_1_708 @runat_27092004_170631_13378 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002s:*
7: Geneseq2003as:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Match	Length	ID	Description
1	2457	100.0	1445	7	Abz224794 Human sol
2	2457	100.0	1856	7	Abz224792 Human sol
3	2457	100.0	2080	4	AA009552 Human tra
4	2448	99.6	1873	7	Abz224793 Human sol
5	2448	99.6	2217	4	AAF55865 Human GLU
6	2162.5	88.0	2087	4	AAF55866 Rat GLUTX
7	2150	87.5	2072	4	AAF55867 Murine GL
8	2123.5	86.4	1954	6	Abx34029 Human can

ALIGNMENTS

RESULT 1	ID	ABZ224794	standard; cDNA; 1445 BP.
ABZ224794	XX	AC	ABZ224794;
XX	AC	ABZ224794;	
DT	07-APR-2003	(first entry)	
XX	Human	solute carrier type 2A nucleic acid.	
DE	Human; solute carrier type 2A; SLC2A; glucose transporter; p53; cancer; cytosolic; gene therapy; gene; ss.		
KW	Human; solute carrier type 2A; SLC2A; glucose transporter; p53; cancer; cytosolic; gene therapy; gene; ss.		
XX	Homo sapiens.		
XX	Key	Location/Qualifiers	
FT	CDS	7..1440	
FT		/*tag= a	
FT		/product= "Human SCL2A"	
XX	XX		
EN	WO200298467-A1.		
XX	12-DEC-2002.		
PD	03-JUN-2002; 2002WO-US017419.		
XX	05-JUN-2001; 2001US-0296076P.		
PR	10-OCT-2001; 2001US-0328605P.		
PR	15-FEB-2002; 2002US-0357253P.		
XX	(EXEL-) EXELJIX INC.		
PA			
XX			

PI Friedmann L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX WPI; 2003-201283/19.
DR P-PSDB; ABP58364.
XX

Identifying candidate p53 pathway modulating agent as therapeutic target
PT for disorders of defective p53 function e.g. cancer, by assaying purified
PT solute carrier family 2 (SLC2A) polypeptide or nucleic acid with a test
PT agent.

XX
PS
PS Disclosure: Page 45; 58pp; English.

XX The present sequence is that of a human solute carrier type 2A (SLC2A)
CC nucleic acid. Genetic screens were designed to identify modifiers of the
CC p53 pathway in *Drosophila* in which p53 was overexpressed. Human
CC orthologues (polynucleotides and polypeptides) of one such modifier were
CC then identified, including the present nucleic acid. SLC2As are glucose
CC transporter proteins with sugar transporter domains. SLC2As nucleic acids
CC and polypeptides are attractive drug targets for the treatment of
CC pathologies associated with a defective p53 signalling pathway, such as
CC cancer. The invention provides *in vitro* and *in vivo* methods of assessing
CC SLC2A function. Modulation of an SLC2A or its binding partners is useful
CC for understanding the association of the p53 pathway and its members in
CC normal and disease conditions and for developing diagnostic and
CC therapeutic modalities of p53-related pathologies. SLC2A-modulating
CC agents that act by inhibiting or enhancing SLC2A expression, directly or
CC indirectly, e.g. by affecting an SLC2A function such as transport or
CC binding activity, can be identified using methods provided. Modulators
CC include small molecules, nucleic acids, antibodies, antisense
CC oligonucleotides and phosphothioate morpholino oligomers (claimed)

XX
SQ Sequence 1445 BP; 182 A; 500 C; 456 G; 307 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.02e-204 Length: 1445
Score: 2457.00 Matches: 477
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-886-954A-1 (1-477) x ABZ24794 (1-1445)

QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
DB 7 ATGACGCCGAGGACCCAGAGGAAACCCAGCCGCTTCTGGGCGCTCTGGCGGCGAGCGG 66
QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
DB 67 CCCCAGCGCGCGCGCTCTCTCGCGCGCTTTCGCGCGCTGCGCGCCACTCAGCTTC 126
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaProPro 60
DB 127 GGCCTTCGCGCTCGGCTACAGCTCCCGCGGCATCCCTAGCTCAGCGCGCGCGCGCG 186
QY 61 AlaProArgLeuAspAlaAlaAlaLaserTyrPheGlyAlaValValThrLeuGlyAla 80
DB 187 GCGCGCGCGCTGAGACGCGCGCGCTCTCGTTCGGGCGCTGTGTGACCTCGGTGCC 246
QY 81 AlaAlaGlyGlyValLeuGlyGlyTyrLeuValAspArgAlaGlyArgGlySerLeu 100
DB 247 GCGCGCGCGGAGTGTGGCGCGCTGGCTGTGTGTGACCGCGCGCGCGCGCGCGCTC 306
QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
DB 307 TTGCTGTGCTCGTGCCTCTGTGGCGCGCTTTCGCGCTCATCCGCGCGCGCGCGCG 366
QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
DB 367 TGGATGCTGCTGGGCGCGCTCTCTCACCAGCGCTGGCGCTGGCGCTGGCTGGCTAG 426
QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
|||||

DB 427 GCCCGGCTTACATCTCCGAAATCGCCTACACAGCAGTCCGGGGGTTGCTGGCTCTCTGT 486
QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGlu 180
DB 487 GTGAGACTAAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 546
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
DB 547 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 606
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
DB 607 TTCATGCCCGAGACCCCGCGCTTCTGTGTGACTCAGCACAGCGCGCGAGAGCCATGGCC 666
QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTyrGluAspProProIleGlyAlaGlu 240
DB 667 GCCCTGGCGGTTCCTGTGGGCTCCAGCAGGCTGGAGACCCCGCCCATCGGGCTGAG 726
QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGly 260
DB 727 CAGAGCTTTCACCTGGCGCTTCTGGCGGAGCCCGGCATCTACAAGCCCTTCATCATCGC 786
QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAlaAlaValMetPheTyrAlaGlu 280
DB 787 GTCTCCCTGATGGCTTCCAGCAGCTGTGGGGGTCAACGGCGTCTGTCTATGCAGAG 846
QY 281 ThrIlePheGluGluAlaLysPheLysSerSerLeuAlaSerValValGlyVal 300
DB 847 ACCATCTTGAAGAGGCCCAAGTTCAAGGACAGCCTGGCGCTCGGTGTGGTGGTGTG 906
QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeu 320
DB 907 ATCCAGGTGCTTTCACAGCTGTGGCGGCTCTCATCATGGACAGAGCGGCGAGGCTG 966
QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
DB 967 CTCCTGGCTTGTGTCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1026
QY 341 LysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
DB 1027 AAGCTGACCCAGGGTGGCGCTGGCAACTCCTCGACGTGGCCATCTCGGCGCTGTCTCT 1086
QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValGlySerMetCysLeu 380
DB 1087 GCACAGCCTGT 1146
QY 381 PheIleAlaGlyPheAlaValGlyTyrGlyProIleProTrpLeuLeuMetSerGluIle 400
DB 1147 TTCATCGCGGCTTTCGCGTGGCTGGGCGCCCATCCCTGGCTCTCTCATGTGCAGAG 1206
QY 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
DB 1207 TTCCTCTGTGATGCAAGGGGTGGGACAGGATCTGGTCTCTCACCACCTGGCTCATG 1266
QY 421 AlaPheLeuValThrLysGluPheSerLeuMetGluValLeuArgProTyrGlyAla 440
DB 1267 GCCTTTCGTGACCAAGAGTTCAGCAGCCTCATGGAGGTCTCTCAGGCCCTATGGAGCC 1326
QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
DB 1327 TTCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1386
QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
DB 1387 GAACCTAAGGAAGACTCTGGAAACAATCAGAGCCCATTTGAGGGGCGA 1437
RESULT 2
ABZ24792
ID ABZ24792 standard; cDNA; 1856 BP.
XX
AC ABZ24792;
XX
DT 07-APR-2003 (first entry)


```

Db      1144 TTATCGCGCGGCTTGGCTGGCTGGGGCCATCCCTCGCTCTCATGTACAGATC 1203
QY      401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
Db      1204 TTCCCTCTGCAATGTCAGGGCGTGCGACAGGCATCTCGTCTCCACCAACTGGTCTCATG 1263
QY      421 AlaPheLeuValThrLysGluPheSerLeuMetGluValLeuArgProTrpGlyAla 440
Db      1264 GCCTTTCGTGACCAAGAGTTACAGACGCTCATGAGGTCCTCAGGCCCTATGGAGCC 1323
QY      441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
Db      1324 TTCTGGCTGGCTCGCTTCGATCTTCAGTCCTTCAGTGCCTTTTCACATTGTTGTCGTCCCT 1383
QY      461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db      1384 GAAACTAAAGGAAGACTCTGGAACAAATACAGGCCCATTTTGGAGGGCGA 1434
RESULT 3
AAD09552
ID      AAD09552 standard; cDNA; 2080 BP.
XX
AC      AAD09552;
XX
DT      10-SEP-2001 (first entry)
XX
DE      Human transporter and ion channel-1 (TRICH-1) cDNA.
XX
KW      Human; transporter and ion channel-1; TRICH-1; vaccine; cystic fibrosis;
KW      gene therapy; amphotrophic lateral sclerosis; amnesia; muscular dystrophy;
KW      hypertension; angina; neurological disorder; asthma; bipolar disorder;
KW      dementia; depression; Alzheimer's disease; epilepsy; mood; arrhythmia;
KW      Pick's disease; ischemic cerebrovascular disease; AIDS; anxiety; stroke;
KW      Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy;
KW      demyelinating disease; mental disorder; Schizophrenia; polymyositis;
KW      muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease;
KW      dermatomyositis; diabetes mellitus; immunological disorder; psoriasis;
KW      rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus;
KW      sickle cell anaemia; Wilson's disease; infertility; Cushing's disease;
KW      scleroderma; pulmonary artery stenosis; nontropic; Addison's disease;
KW      malabsorption syndrome; hypercholesterolaemia; cancer; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      CDS
FT      46..1479
FT      /tag= a
FT      /product= "Human TRICH-1 protein"
XX
PN      WO200146258-A2.
XX
PD      28-JUN-2001.
XX
PF      22-DEC-2000; 2000WO-US035095.
XX
PR      23-DEC-1999; 99US-0172000P.
PR      14-JAN-2000; 2000US-0176083P.
PR      21-JAN-2000; 2000US-0177332P.
PR      28-JAN-2000; 2000US-0178572P.
PR      02-FEB-2000; 2000US-0179758P.
PR      10-FEB-2000; 2000US-0181625P.
XX
PA      (INCY-) INCYTE GENOMICS INC.
XX
PI      Baughn MR, Burford N, Au-Young J, Lu DAM, Yang J, Reddy R;
PI      Lal P, Hillman JL, Azimzai Y, Yue H, Nguyen DB, Yao MG, Gandhi AR;
PI      Tang YT, Khan FA;
XX
DR      WPI; 2001-418042/44.
DR      P-PSDB; AA504888.
XX
PT      Novel human transporter and ion channel proteins useful for treating and

```

```

PT      preventing transport, neurological, muscle and immunological disorders.
XX
PS      Claim 5; Page 140-141; 160pp; English.
XX
The present sequence is transporter and ion channel-1 (TRICH-1) cDNA.
TRICH is used as vaccine. TRICH is useful for treating a disease or
condition associated with decreased expression of functional TRICH, such
as transport disorder including amphotrophic lateral sclerosis, cystic
fibrosis, Becker's muscular dystrophy, Charcot-Marie Tooth disease,
Duchenne muscular dystrophy, angina and hypertension, neurological
disorders including Alzheimer's disease, amnesia, bipolar disorder,
dementia, depression, epilepsy, ischemic cerebrovascular disease,
stroke, cerebral neoplasms, Pick's disease, Huntington's disease and
Parkinson's disease, demyelinating diseases, mental disorders including
mood, anxiety, Schizophrenia and seasonal affective disorder, muscle
disorder including cardiomyopathy, myocarditis, polymyositis,
dermatomyositis, arrhythmias and asthma and immunological disorders
including AIDS, adult respiratory distress syndrome (ARDS), allergies,
anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's
syndrome, systemic lupus erythematosus and other diseases including
sickle cell anaemia, Wilson's disease, cataracts, infertility, pulmonary
artery stenosis, Grave's disease, Cushing's disease, Addison's disease,
glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers,
psoriasis and viral, bacterial, fungal, helminthic and protozoal
infections. TRICH DNA is useful in gene therapy and in diagnostic
purposes
XX
SQ      Sequence 2080 BP; 300 A; 688 C; 632 G; 460 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      1.65e-204      Length:      2080
Score:          2457.00      Matches:      477
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              4      Gaps:      0

US-09-886-954A-1 (1-477) x AAD09552 (1-2080)
QY      1 MetThrProGluAspProGluGlnThrGlnProLeuLeuGlyProProGlySerAla 20
Db      46 ATGACGCCCGAGGACCCAGAGAAACCCAGCGCTTCTGGGGCCCTCTGCGCGCAGCGG 105
QY      21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
Db      106 CCCCGCGCGCGCGGCTTCTCTCGCGCCTTGGCGGCTGCCCTGGGCCCATCTAGCTTC 165
QY      41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
Db      166 GGCCTTCGCGCTCGGTACAGCTCCCGGCCATCCCTAGCTGACGCGCGCGGCCCGCG 225
QY      61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
Db      226 GCCCGCGCCCTGGACGACGCGCGCCCTCTCGTTCGGGGCTGTGCTGACCTGGGTGCC 285
QY      81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db      286 GCGGCGGGGGAGTGTCTGGCGGCTGGCTGGTGGACCGCGCGCGGCGAAGCTGAGCCTC 345
QY      101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db      346 TTGCTGTGCTCGTCCCTCTCGTGGCGGCTTTCGCGTCATCACCGCGCGCGCCAGACGTG 405
QY      121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db      406 TGGATGCTGTGGGGGGCGGCTCTCTACCGGCCCTGGGCTGCGGTGTGGCTCCCTAGTGTG 465
QY      141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
Db      466 GCCCGCGTCTACATCTCCGAAATCGCTACCCAGCAGTCCGGGGGTGCTCGGCTCCTGT 525
QY      161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180

```



```
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAalaProPro 60
DB 124 GGCCTTCGGGCTCGGCTACAGCTCCCGGGCCATCCCTAGCTGACGGCGCCGCCCGG 183
QY 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
DB 184 GCGCCGCGCTCGACGACGCGCGCCCTCCTGGTTGGGGCTGTCTGTGACCTCGGTGCC 243
QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
DB 244 GCGCGGGGGAGTCTCGGGGGCTGGCTGGTGGACCGCGCGGGCGAAGCTGAGCCCTC 303
QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
DB 304 TTGCTGTGCTCGTGCCCTTCGTGGCGGCTTTGGCGTCATCACCGCGGCCGAGACGTG 363
QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
DB 364 TGGATGCTGCTGGGGGGCGCCCTCCTCACCGGCTGGCGTGGGTGGTTCCTCCCTAGTG 423
QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
DB 424 GCGCGGGCTCATCATCTCGAAATCGCCTACCCAGCAGTCCGGGGGTGCTCGGCTCCCTGT 483
QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
DB 484 GTGCAGCTAATGTCGTCTGGCATCTCTGGCTACCTGGCAGGCTGGGTGGTGAG 543
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
DB 544 TGGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 603
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
DB 604 TTCATGCCCGAGACCCCGGCTTCCTGCTGACTCAGCACAGCGCCAGGAGCCATGGCC 663
QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProPheGlyAlaGlu 240
DB 664 GGCCTCGGTTCTCTGGGGCTCCGAGCAGGCTGGGAGACCCCCCATCGGGGCTGAG 723
QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleLeGly 260
DB 724 CAGAGCTTTCACCTGGCCCTGTGGCGGACGCCGCGCATCTACAAGCCCTTCATCATCGC 783
QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
DB 784 GTCCTCCCTGATGGCTTCAGCAGCTGTCGGGGGTCAAGCCGTATGTTCTATGCAGAG 843
QY 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
DB 844 ACCATCTTTGAAGAGGCCAAGTTCAAGGACACAGCCTGGCCCTCGGTCTCGTGGGTGTC 903
QY 301 IleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgArgLeu 320
DB 904 ATCCAGGTGCTGTTACAGCTGTGGGGCTCTCATCATGACAGAGCGGGCGGAGGCTG 963
QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
DB 964 CTCCTGGTCTGTAGGTGTGTATGTGTTTCAGCAGAGTGGCTTGGCGCCCTACTTTC 1023
QY 341 LysLeuThrGlnGlyProGlyAsnSerHisValAlaIleSerAlaProValSer 360
DB 1024 AAGCTGACCCAGGGTGGCCCTGGCAACTCCTCGCAGGTGGCCATCTCGCGGCTGTCTCT 1083
QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
DB 1084 GCACAGCCTGTGTATGCCAGCTGGGGCTGGGCTGGCTGGCTGGCGGCAACATGCGCCTC 1143
QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuMetSerGluIle 400
DB 1144 TTCATCGCGGCTTTGGGTGGCTGGGGGCCCATCCCTCGCTCTCATGTACAGATC 1203
QY 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
```

```
DB 1204 TTCCTCTGCATGTCAAGGGGTGGCGACAGCATCTGGTCTCCACCACTGGCTCATG 1263
QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
DB 1264 GCTTTTCTCGTGACCAAGGAGTTTCAGACAGCCTCATGAGGTCTCTCAGGCCCTATGAGCC 1323
QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
DB 1324 TTCGTGGCTGGCTCGGCTTTCGCACTCTTCAGTGTCTTTCACCTTGTCTGTGTCCT 1383
QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
DB 1384 GAAATTAAGGAAGAGACTCTGGAACAAATCACAGCCCATTTTGAGGGGCGA 1434

RESULT 5
AAF55865
ID AAF55865 standard; cDNA; 2217 BP.
XX
AC AAF55865;
XX
DT 17-APR-2001 (first entry)
XX
DE Human GLUTX1 coding sequence.
XX
KW Human; GLUTX; gene therapy; vaccine; hexose transport modulator;
KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS Homo sapiens.
XX
PN WO200104145-A2.
XX
PD 18-JAN-2001.
XX
PF 14-JUL-2000; 2000WO-IB001042.
XX
PR 14-JUL-1999; 99US-0143907P.
PR 27-AUG-1999; 99US-0151140P.
PR 23-FEB-2000; 2000US-0184285P.
PR 13-JUL-2000; 2000US-00616132.
XX
PA (UYLA-) UNIV LAUSANNE.
XX
PI Thorens B, Ibberson M, Uldry M;
XX
DR WPI; 2001-112615/12.
DR P-PSDB; AAB66932.
XX
Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
prevention, diagnosis and treatment of hexose transport disorders, e.g.
ischemia and diabetes.
XX
PS Claim 3; Page 70-71; 124pp; English.
XX
The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
AAB66932-AAB66941). The GLUTX proteins are related to the facultative
glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
function. The GLUTX proteins may be used in the diagnosis, prevention and
treatment of hexose transport disorders such as ischaemia, diabetes,
hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
neurodegenerative disease. The present sequence is the coding sequence
for human GLUTX1
XX
SQ Sequence 2217 BP; 320 A; 750 C; 672 G; 475 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,1e-203 Length: 2217
Score: 2448.00 Matches: 475
Percent Similarity: 99.73% Conservativeness: 1
Best Local Similarity: 99.58% Mismatches: 1
Query Match: 99.63% Indels: 0
DB: 4 Gaps: 0
```


CC for rat GLUTX1

XX Sequence 2087 BP; 355 A; 673 C; 580 G; 479 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8,37e-179 Length: 2087
 Score: 2162.50 Matches: 408
 Percent Similarity: 92.26% Conservatives: 33
 Best Local Similarity: 85.36% Mismatches: 36
 Query Match: 88.01% Indels: 1
 DB: 4 Gaps: 1

US-09-886-954A-1 (1-477) x AAF55867 (1-2087)

QY 1 MetThrProGluAspProGluGlnThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 |||
 DB 30 ATGTGCGCGAGAGCCCGAGAAAACACAGCGCGCTGTTCGGTTCGCGCGCGCCAGGGCT 89
 |||
 QY 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 |||
 DB 90 CCGGGGGCGCGGGCTTCCTTGCACCTTCGCGCGCGCTGGGCCCGCTCAGCTTC 149
 |||
 QY 41 GlyPheAlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 |||
 DB 150 GCGTTCGCGCTTGGCTACAGCTCCCGGCATCCCGAGCTCGCGCGCACCCCGCT 209
 |||
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 |||
 DB 210 GCGCTACGCTCGGAGACATCGCGCTCTCGTTCGGCGCGCTGTGACCTCGCGCGCT 269
 |||
 QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlyLeuSerLeu 100
 |||
 DB 270 GCGGAGGGGGCGTCTGGCGCGCTGGCTTCCTGGACCGTTCGAGGGCGCAGCTGACCTC 329
 |||
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 |||
 DB 330 CTGCTCTGACCGTTCGCTGACCGCTTTTGTGTCTCATCACCGCGCGCGGATGTG 389
 |||
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 |||
 DB 390 TGGATGCTCTCGGAGCGCGCTCTCATCCGCGCTAGCGTTCGGAGTCCGCTCATTAGTG 449
 |||
 QY 141 AlaProValTyrIleSerGluLeuAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 |||
 DB 450 GCACCGGTCTATATCTCGGAATCGCTACCCAGCGCTTCGAGGACTGCTGGCTCCTGT 509
 |||
 QY 161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 |||
 DB 510 GTGCAGCTGATGTTGTCTACTGGCATCTCTTGGCTACGTGGCAGGCTGGGTCTTAGAA 569
 |||
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 |||
 DB 570 TGGCGTGTGTGGCTGTGTGGCTGTGTGCCCCCACCCTCATGTGCTGTCTATGTGC 629
 |||
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgGlnGlnAlaMetAla 220
 |||
 DB 630 TATATGCCCGAGNCCCGAGCTTCTCTTGACGCAACACACAGTACCAGGAGCCATGGCT 689
 |||
 QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
 |||
 DB 690 GCGCTCGCTCTCTGTGGGCTCTGAGGAGGGCTGGGAAAGAGCCCTGTGGGGCTGAG 749
 |||
 QY 241 ---GlnSerPheHisLeuAlaLeuArgGlnProGlyIleTyrIlePhePheIleIle 259
 |||
 DB 750 CACCGAGGCTTCAGCTGGCTATGTGAGCGCGCTTGGTGTCCAGAGCCCTCATCATC 809
 |||
 QY 260 GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla 279
 |||
 DB 810 GGCATTGGCTCATGTCTTCAGCAGCTGTCHAGGGGTCAAGCCCATCATGTTCTATGCC 869
 |||
 QY 280 GluThrIlePheGluGlnAlaLysPheLysAspSerSerLeuAlaSerValValValGly 299
 |||
 DB 870 AACACCATCTTTGAGGAGGCAAGTTCAAGGACAGAGCGCTGGCCTCGGTCACTGTGGGC 929
 |||

QY 300 ValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgArg 319
 ::
 DB 930 ATCATCCAGGTTCTGTTCACTGCTGTGGCGGCTTCATCGACAGAGAGGGCGAAAA 989
 ::
 QY 320 LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr 339
 |||
 DB 990 CTGCTTCTGGCTTGTGGGTGTGATCATGTTTTCAGCATGAGCGCTTTGGTACCTAC 1049
 |||
 QY 340 PheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProVal 359
 |||
 DB 1050 TTCAAACTGACCCAGAGTGGCCCGACCACTCTCCCATGTAGGCTCTCTGTGCCCCATC 1109
 |||
 QY 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCys 379
 |||
 DB 1110 TCCGACAGAGCTGCTGATGTTCCACTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGC 1169
 |||
 QY 380 LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGlu 399
 |||
 DB 1170 CTCTTCATCGCTGTTTTCAGTAGGCTGGGACCCATCCCTGGCTCTCTCATGTCCAGAG 1229
 |||
 QY 400 IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeu 419
 |||
 DB 1230 ATCTTCCCTCTGCATCAAGGGTGTGGCTACCGGCTGTGTGTCTCACCACCTGGTTC 1289
 |||
 QY 420 MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGly 439
 |||
 DB 1290 ATGGCTTCTTGTGACCAAGAGTTTACAGCATCTGGAGATCTCTCAGACCTTACGGC 1349
 |||
 QY 440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
 |||
 DB 1350 GCCTTCTGGCTCACCGCTGCTTCTGTATCTCTCAGCTGCTCTTTCACGCTCACCTTGTTC 1409
 |||
 QY 460 ProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 |||
 DB 1410 CCTGAGACTAAAGGAGGAGGACTCTGGAACAAATCAGACCCCATTTTCAGGGGACGG 1463
 |||

RESULT 7

AAF55867
 ID AAF55867 standard; cDNA; 2072 BP.

XX AAF55867;

DT 17-APR-2001 (first entry)

XX Murine GLUTX1 coding sequence.

XX Murine; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.

XX Mus sp.

XX WO200104145-A2.

XX 18-JAN-2001.

XX 14-JUL-2000; 2000WO-IB001042.

XX 14-JUL-1999; 99US-0143907P.

PR 27-AUG-1999; 99US-0151140P.

PR 23-FEB-2000; 2000US-0184285P.

PR 13-JUL-2000; 2000US-00616132.

XX (UYLA-) UNIV LAUSANNE.

XX Thorens B, Iberson M, Uldry M;

XX WPI; 2001-112615/12.

DR P-PSDB; AAB66934.

XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
 PT prevention, diagnosis and treatment of hexose transport disorders, e.g.

PT ischemia and diabetes.

PS Claim 3; Page 73-74; 124pp; English.

XX
CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and AAB6932-AAB6941). The GLUTX proteins are related to the facultative glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport function. The GLUTX proteins may be used in the diagnosis, prevention and treatment of hexose transport disorders such as ischaemia, diabetes, hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a neurodegenerative disease. The present sequence is the coding sequence for murine GLUTX1

SQ Sequence 2072 BP; 351 A; 673 C; 584 G; 464 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,02e-177	Length:	2072
Score:	2150.00	Matches:	409
Percent Similarity:	92.47%	Conservative:	33
Best Local Similarity:	85.56%	Mismatches:	34
Query Match:	87.51%	Indels:	2
DB:	4	Gaps:	2

US-09-886-954A-1 (1-477) x AAF55867 (1-2072)

QY	1	MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla	20
DB	21	ATGTCCTCCGAGGACCCAGAGACGACGCGCTATTGCGGCCACCGAAGCCAGGACT	80
QY	21	ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuLeuGlyProLeuSerPhe	40
DB	81	CCCCGGCCCGCGGCTCTCGCTTCCTTCGCGCGCGCTCTGGGACCCCTCAACTTC	140
QY	41	GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuLeuGlnArgAlaAlaPro	60
DB	141	GGCTTCGGCGCTCGGCTACAGCTCCCGCCCATCCCGACCTCGCGCGCACCGCCCG	200
QY	61	AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValAlaValThrLeuGlyAla	80
DB	201	GCCCTGCGCTCGGAGACAATGCGGCTCTCGGTTGCGGGCGCTCGTACCTCGGCGCT	260
QY	81	AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlyLeuSerLeu	100
DB	261	GCTCAGGGGCGATCACTGGCGGCTGGCTCTGGACCGTTCCAGGCGCAAGCTGAGCCTC	320
QY	101	LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal	120
DB	321	TTGCTCTGCACCGTCCCTTCGTGACTGGCTTTGCTGTCATCACGCGCGCGCGGATGTG	380
QY	121	TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal	140
DB	381	TGGATGCTGCTCGGAGCGCGCTCTCTACCGGCGCTTGCTCGGAGTCCGCTCACTAGTG	440
QY	141	AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys	160
DB	441	GCACCGGTTTACATCTCGAAATCGCTTACCAGCTGTCCGAGGACTGCTCGGCTCCTGT	500
QY	161	ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu	180
DB	501	GTGACGTGATGGTGTGTCATGGCATCTCTCGGCTATGTGGCAGGCTGGTCTCTAGAG	560
QY	181	TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys	200
DB	561	TGGCGCTGGCTGGCGCTGCTGGCTGTGTGCCCGCCACCCCTCATGCTGCTCATGTC	620
QY	201	PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla	220
DB	621	TACATGCCGAGACCCACGTTTCTCCTCACTCAACACCACTACAGGAGGCCATGGCT	680
QY	221	AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGlu	240
DB	681	GCCTTGGCGCTTCTGTGGGGCTCTGAGAGGGCTGGGAAGAGCCCGCTGTGGGGCTGAG	740

QY	241	---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIysProPheIleIle	259
DB	741	CACCAAGGCTTCCAGCTGGCCCTGCTGAGGGCCCTTGGCATCTCAAGCCCTCATATC	800
QY	260	GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla	279
DB	801	GGCATTTCCCTCATGGTCTTCACAGAGCTGTTCAGGGGTCAATGCTATCATGTTCTATGCC	860
QY	280	GluThrIlePheGluGluAlaIlyPheIysAspSerSerLeuAlaSerValValGly	299
DB	861	AACAGCATCTTCGAGGAGGCCAAGTTCAAGACACAGCAGCTGGCTCGTCACTGTGGGC	920
QY	300	ValIleGlnValLeuPheThrAlaValAlaAlaIleMetAspArgAlaGlyArg	319
DB	921	ATAATCCAGGTCCTGTTCACTGCTGTGGCGCCCTCATCATGGACAGAGGGCGAAGG	980
QY	320	LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr	339
DB	981	CTGCTCTCTGGCTTGTTCGGGTGTGATCATGTGTTCAGTATGAGTGCCTTTGGTACCTAC	1040
QY	340	PheIysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProVal	359
DB	1041	TTCAAACTGACCCAGAGCTCCCCAGCAACTCTCTCCCACTAGGCTG---GTGCCCATC	1097
QY	360	SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCys	379
DB	1098	CGCGGAGAGCTGTGGATGTCCAAGTGGACTGGCTGGCTGGCTGTAGGCGAGCATGTGC	1157
QY	380	LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGlu	399
DB	1158	CTCTTCATGCTGGCTTTCGGTGGCTGGGAGCCCATCCCTGGCTCTCATGTGCAGAG	1217
QY	400	IlePheProLeuHisValIysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeu	419
DB	1218	ATCTTCCCTCTGCATGTCAAGGTTGGCTACCGGCATCTGTCTCTCACCACCTGGTTC	1277
QY	420	MetAlaPheLeuValThrLysGlyPheSerSerLeuMetGluValLeuArgProTyrGly	439
DB	1278	ATGGCTTTCTAGTGACCAAGAGTTCAACAGCGTCATGGAGATGCTCAGACCTACGGT	1337
QY	440	AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal	459
DB	1338	GCCTTCTGGCTCACCGCTGCCTTCTGGCTCTCAGTGTCTTTCACACTGACCGTTGTC	1397
QY	460	ProGluThrIysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg	477
DB	1398	CCTGAGACTAAGGAGGAGCTCTGGAACAAGTCACAGCCCATTTTCGAGGGACGA	1451

RESULT 8
ABX34029
ID ABX34029 standard; cDNA; 1954 BP.
XX
AC ABX34029;
XX
DT 10-FEB-2003 (first entry)
XX
DE Human cancer suppressing protein PP7425.
XX
KW Human; gene; ss; cancer suppressing protein; cancer.
XX
OS Homo sapiens.
XX
FN CN1351081-A.
XX
PD 29-MAY-2002.
XX
PF 31-OCT-2000; 2000CN-00127102.
XX
PR 31-OCT-2000; 2000CN-00127102.
XX
PA (SHAN-) SHANGHAI INST ONCOLOGY.
PI Gu J;

XX 18-MAY-2001; 2001WO-US016450.
 XX 19-MAY-2000; 2000US-0205515P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Birse CE, Rosen CA;
 XX WPI: 2002-122018/16.
 XX P-PSDB; ABB89717.
 XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and prevention of neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal and proliferative disorders.
 XX Claim 4; SEQ ID NO 688; 2081pp + Sequence Listing; English.
 XX The invention relates to novel genes (ABL89449-ABL90853) and proteins (ABB89040-ABB90444) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 1156 BP; 177 A; 362 C; 345 G; 255 T; 0 U; 17 Other;
 Alignment Scores:
 Pred. No.: 5,45e-137 Length: 1156
 Score: 1679.50 Matches: 344
 Percent Similarity: 72.12% Conservative: 0
 Best Local Similarity: 72.12% Mismatches: 10
 Query Match: 68.36% Indels: 125
 DB: 6 Gaps: 1
 US-09-886-954A-1 (1-477) x ABL90126 (1-1156)
 QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 DB 53 ATGACGCCGAGGACCCAGAGAAACCCAGCCGCTTCCTGGGCGCTTCCTGGCGGAR---- 108
 QY 21 ProArgGlyArgArgValPheLeuAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 DB 108 ----- 108
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 DB 108 ----- 108
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 DB 108 ----- 108
 QY 81 AlaAlaGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 DB 108 ----- 108
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 DB 108 ----- 108

QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 DB 108 ----- 108
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 DB 109 -----GTCTACATCTCCGAATCCCTACCAGCAGTCCGGGGTGTCTGGCTCTCTGT 162
 QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 DB 163 GTGCAGCTAATGGTCGTCTCGGCATCTCTGGCTACCTGGCAGGCTGGTGTGGAG 222
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 DB 223 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 282
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAla 220
 DB 283 TTCATGCCCGAGACCCCGGCTTCCTGCTGACTCAGCAGCGCGCAGGAGGCCATGGCC 342
 QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
 DB 343 GGCCTGCGNWTTCCTGTGGGCTCCGAGCAGGGCTGGGAAGACCCCCCATCGGNGCTGAG 402
 QY 241 GlnSerPheHisLeuAlaLeuArgGlnProGlyIleTyrLysProPheIleIleGly 260
 DB 403 CAGAGCTTTCACCTGGCCCTGCTGGGCANCCCGGCATCTACAGCCCTTCATCATCGGC 462
 QY 261 ValSerLeuMetAlaPheGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
 DB 463 GTCTCCCTGATGGCTTCACAGCTGTCGGGGGTCAACGGCTCATGTTCTATGACAGAG 522
 QY 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
 DB 523 ACCATCTTTGAAGAGGCCAAGTTCAAGGACAGCAGCCCTGGCTGGTGGTGGTGGT 582
 QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeu 320
 DB 583 ATCCAGGTGCTGTTCACAGCTGTGGCGCTCTCATCATGACAGCAGGCGGAGGCTG 642
 QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
 DB 643 CTCTCTGGTCTTGTGAGGT 702
 QY 341 LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
 DB 703 AAGCTGACCCAGGCTGGCCCTGGCAACTCTCTCGCAGCTGGCATCTCGGCGCTGTCT 762
 QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
 DB 763 GCACAGCCTGTGTGATGCCAGCGTGGGGCTNGCCCTGGCTGGCTGGCGAGCATGTGCT 822
 QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400
 DB 823 TTCATCGCCGCTTGTGGTGGCTGGGGGCCATCCCTGGCTCTCTCATGTACAGATC 882
 QY 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuLeuThrAsnTrpLeuMet 420
 DB 883 TTCCTTGTGATGTCAAGGGCGTGGCAGAGCAWYTGCGTCTCCAMCAACTGGGCTCATG 942
 QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuLeuArgProTyrGlyAla 440
 DB 943 GSCCTTCTYGGACCAAGGAGTTCAGCAGCCCTCATGGAGGTCTCTCAGGCCCTATGGAG 1002
 QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 DB 1003 TTCCTGGCTTGGCTCCGCTTCTGTGATCTTGTGATCTTGTGATCTTGTGATCTTGTG 1061
 QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 DB 1062 GAAACTTAAGGAAAGACTCTGGAACAAATCACAGCCCATTTTGGAGGGGCGA 1112
 RESULT 10

AAD12574
 ID AAD12574 standard; cDNA; 1461 BP.
 XX
 AC AAD12574;
 XX
 DT 25-SEP-2001 (first entry)
 XX
 DE Human protein having hydrophobic domain encoding cDNA clone HP10784.
 XX
 KW Human; hydrophobic domain; gene therapy; nutritional supplement;
 KW cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;
 KW multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;
 KW haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;
 KW Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;
 KW haemostatic; thrombolytic; tumour growth inhibitor; anabolic;
 KW contraceptive; antiinfectivity; antiinflammatory; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 61..849
 FT /*tag= a
 FT /product= "Human protein having hydrophobic domain"
 FT /note= "CDS is specifically claimed in claim 3"
 XX
 PN WO200149728-A2.
 XX
 PD 12-JUL-2001.
 XX
 PF 28-DEC-2000; 2000WO-JP009359.
 XX
 PR 06-JAN-2000; 2000JP-00000585.
 PR 06-JAN-2000; 2000JP-00000588.
 PR 11-JAN-2000; 2000JP-00002299.
 PR 03-FEB-2000; 2000JP-00026862.
 PR 03-MAR-2000; 2000JP-00058367.
 XX
 PA (PROT-) PROTEGENE INC.
 PA (SAGA) SAGAMI CHEM RES CENT.
 XX
 PI Kato S, Kimura T;
 XX
 DR WPI; 2001-418355/44.
 DR P-PSDB; AAE06579.
 XX
 PT Human proteins with hydrophobic domains and the nucleic acids encoding
 PT them, useful for preventing diagnosing and treating e.g. cancer,
 PT Alzheimer's and inflammation.
 XX
 PS Claim 4; Page 287-289; 563pp; English.
 XX
 CC The present sequence is human protein with hydrophobic domain encoding
 CC cDNA clone HP10784. The polynucleotide and polypeptide of the invention
 CC may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate polypeptide expression. The polynucleotides
 CC may be used to produce the polypeptide, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the protein. The
 CC polynucleotides and its complementary sequences may also be used as DNA
 CC probes in diagnostic assays and also used in gene therapy. The
 CC polypeptides may also be used as antigens in the production of antibodies
 CC and in assays to identify modulators of polypeptide expression and
 CC activity. The polypeptides and nucleic acids may be used as nutritional
 CC supplements, to modulate cytokine and cell proliferation activity, to
 CC modulate immune stimulation or suppression (e.g. for the treatment of
 CC microbial infections and autoimmune disorders such as multiple sclerosis,
 CC rheumatoid arthritis and insulin-dependent diabetes), to modulate
 CC haematopoiesis, to modulate tissue growth activity (e.g. for the
 CC treatment of Parkinson's disease, Huntington's disease and Alzheimer's
 CC disease), to modulate activin and inhibin activity (e.g. for controlling
 CC fertility), to modulate chemotactic and chemokinetic activity, to
 CC modulate haemostatic and thrombolytic activity, to modulate receptor
 CC ligand activity, to modulate inflammation and to inhibit tumour growth

SQ Sequence 1461 BP; 195 A; 501 C; 456 G; 309 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 6,03e-126 Length: 1461
 Score: 1554.50 Matches: 324
 Percent Similarity: 67.85% Conservative: 1
 Best Local Similarity: 67.64% Mismatches: 1
 Query Match: 63.27% Indels: 153
 DB: 4 Gaps: 2

US-09-886-954A-1 (1-477) x AAD12574 (1-1461)

QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 DB 61 ATGAGCGCCGAGGACCCAGAGGAACCCAGCGCTTCTGGGGCTCTCTGGCGGACGCG 120
 QY 21 ProArgGlyArgArgValPheLeuAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 DB 121 CCCCAGCGCGCGCGCTTCTCTCGCGCTTCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 180
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 DB 181 GGCCTTCGCGCTCGGCTACAGCTCCCGGCATCTCTAGCTTCAGCGCGCGCGCGCGCGCG 240
 QY 61 AlaProArgLeuAspAspAlaAlaAlaSerTTPheGlyAlaValValThrLeuGlyAla 80
 DB 241 GCCCGCGCGCTGGACGACGCGCGCGCTCTCTGGTTGCGGGCTGCTGTAACCTGCGTGC 300
 QY 81 AlaAlaGlyGlyValLeuGlyGlyTyrProValAspArgAlaGlyArgGlySerLeu 100
 DB 301 GCGCGCGCGGAGTGTCTGGCGCGCTGGTGTGTGGACCGCGCGCGCGCGCGCGCGCGCTC 360
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 DB 361 TTGCTGTGCTCCGTCGCCCTTCGTGCGCGCTTTGCGCTCATCACCGCGCGCGCGCGCG 420
 QY 121 TrpMetLeuLeuGlyGlyArgLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 DB 421 TGGATGCTGCTGGGGGCGCGCTCTCTCACCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 480
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 DB 481 GCCCGCGCTACATCTCCGAAATCGCTACCCAGCAGTCCCGGGGTGCTCGGCTCTCTCT 540
 QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGlu 180
 DB 541 GTGAGCTAATGTCGTGTCGCGCATCTCTCTGCGCTACTCTGCGAGGCTGGGTGCTGGAG 600
 QY 181 TrpArgTyrLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 DB 601 TGGCGCTGGCTGGCTGTGCTGGGCTGCGTGGCGCGCTCTCTCTCTCTCTCTCTCTCTCT 660
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAla-Metal 220
 DB 661 TTCATGCCGAGACCCCGCGCTTCTCTGCTGACTACACAGCGCGCGCGCGCGCGCTGCTC 718
 QY 220 aAlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaG 240
 DB 718 ----- 718
 QY 240 uGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleG 260
 DB 718 ----- 718
 QY 260 yValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaG 280
 DB 718 ----- 718
 QY 280 uThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
 DB 718 ----- 718
 QY 300 lIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLe 320

```

Db 718 ----- 718
QY 320 uLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPh 340
Db 719 -----CTGGTCTTGTGTCAGGTGTGTGTCATGTGTTCAGCAGAGTGTCTTCGGCGCTACTT 774
QY 340 eLysLeuThrGlnGlyGlyProGlyValSerHisValAlaIleSerAlaProValSe 360
Db 775 CAAGCTGACCCAGGTGGCCCTGGCAACTCTCGCAGGTGGCCATCTCGGCGCCTGTCTC 834
QY 360 zLaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLe 380
Db 835 TGACAGAGCTGTGTATGCCAGCGTGGGGCTGGCTGCTGGCTGGCGGTGGCAGATGTGCT 894
QY 380 uPheIleAlaGly-PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluI 400
Db 895 CTTCATCCCGGAGTCT----- 911
QY 400 lePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuM 420
Db 911 ----- 911
QY 420 etAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyA 440
Db 912 -----CTCAGGCCCTATGGAG 927
QY 440 laPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValP 460
Db 928 CCTCTGGCTTGCCTCCCGCTTCTGCATCTTCAGTGTCTCTTTCATCTTGTCTGTCTCC 987
QY 460 roGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 988 CTGAACCTAAGGAAGACTCTGGAAACAAATCACGCCATTTTGGGGGGCA 1040

RESULT 11
AAF55871
ID AAF55871 standard; cDNA; 2011 BP.
XX
AC AAF55871;
XX
DT 17-APR-2001 (first entry)
XX
DE Rat GLUTX3 coding sequence.
XX
KW Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;
KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS Rattus sp.
XX
PN W0200104145-A2.
XX
PD 18-JAN-2001.
XX
PF 14-JUL-2000; 2000WO-1B001042.
XX
PR 14-JUL-1999; 39US-0143907P.
PR 27-AUG-1999; 39US-0151140P.
PR 23-FEB-2000; 2000US-0184285P.
PR 13-JUL-2000; 2000US-00616132.
XX
PA (UCLA-) UNIV LAUSANNE.
XX
PI Thorens B, Ibberson M, Uldry M;
XX
DR WPI; 2001-112615/12.
XX
DR P-PSDB; AAB66938.
XX
PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
PT prevention, diagnosis and treatment of hexose transport disorders, e.g.
PT ischemia and diabetes.
XX

```

PS Claim 3; Page 82-83; 124pp; English.

XX The present invention relates to GLUTX proteins (AAF55865-AAF55871 and AAB66932-AAB66941). The GLUTX proteins are related to the facultative glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport function. The GLUTX proteins may be used in the diagnosis, prevention and treatment of hexose transport disorders such as ischaemia, diabetes, hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a neurodegenerative disease. The present sequence is the coding sequence for rat GLUTX3

SQ Sequence 2011 BP; 337 A; 621 C; 593 G; 459 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 5,75e-75 Length: 2011
 Score: 972.50 Matches: 219
 Percent Similarity: 61.19% Conservative: 79
 Best Local Similarity: 44.97% Mismatches: 169
 Query Match: 39.58% Indels: 21
 DB: 4 Gaps: 7

US-09-886-954A-1 (1-477) x AAF55871 (1-2011)

```

QY 3 ProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAlaProArg 22
Db 64 CCGCGTCGCTGAAGAG-----AAGGCACGGCGCGGGCGCTTGCAA 105
QY 23 GlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPhe 42
Db 106 AACAGAAGGGTTCCTGGCCACCTTCGCTCCCGTGTGGCAATTCAGCTTTGGGTAT 165
QY 43 AlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaPro 62
Db 166 GCCCTGGTCTACATCTCCAGTCATCCCTCGCTGAAGCGCTCTTCTGACCCAGCACTA 225
QY 63 ArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAla 82
Db 226 CGCTGGACAAAATCCAGGCATCTCTGGTTCGGTCCGCTGTTACCCCTGGTCCCGCGCT 285
QY 83 GlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuLeu 102
Db 286 GGGGCGCTCAGTGCTATGCTACTCAATGACCTCTTGGCCCGAAGCTCAGCATCATGTTT 345
QY 103 CysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
Db 346 TCCGCTGTCCCTCGGCATTTGGCTATGCACTCATGGCTGGTCCCGTGGCTCTGGATG 405
QY 123 LeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaAlaSerLeuValAlaPro 142
Db 406 CTTCTGTGGGAGGATGCTGACAGGCTTTTGGCGGGGACTCACTGCTGCTGCATCCCG 465
QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
Db 466 GTGTACGTGCTGAGATTGCACCCCGGGTGTTCGCGGGGCCCTGGGGGCCACGCCGAG 525
QY 163 LeuMetValValValGlyIleLeuLeuAlaTyrTrpLeuAlaGlyTrpValLeuGluTrpArg 182
Db 526 CTCATGGCCGTGTTTGGATCCCTGCTCTATGCCCTTGGTCTTCTGCTGCTTGGCGA 585
QY 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuLeuMetCysPheMet 202
Db 586 TGGCTTGGCGTGGCGGGGAGGGCGCTGTTCTGTCATGATCTCCTGCTGTCTAGCTTCA 645
QY 203 ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaAlaLeu 222
Db 646 CCCAACTCGCTTCCTTCTACTGTCTAAGAGCGGGATGAGGAAGGACTGCAGCGCGCTG 705
QY 223 ArgPheLeuTrp-----GlySerGluGlnGlyTrpGluAspProProIleGly--- 238
Db 706 -----ATCTGGTTCGAGCGGACTCTGAGTCCACTGGGAGTTTGAGCAGATCCAGGAC 759
QY 239 -----AlaGluGlnSerPheHisLeuAlaLeuLeuArg-----GlnProGlyIleTyr 254
Db 239 -----GlnProGlyIleTyr-----GlnProGlyIleTyr-----GlnProGlyIleTyr

```

```
Db 760 AACGTGGGAGACAGTAGCCAGTGTCTGTGGGGAGGCGCTGGAGCCCGTGTGTAC 819
QY 255 LysProPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAla 274
Db 820 CGCCCTATTCTCATACAGTGTGTATGGCTTTCTGCAGCAGCTGCACAGGCATCTCCC 879
QY 275 ValMetPheTyrAlaGluThrIlePheGluGluAlaLys---PheLysAspSerSerLeu 293
Db 880 ATCCCTCGTGACTACAGACCACTTCGACAGCAGCTCGGTGCTGCTCCAGAG 939
QY 294 AlaserValValIleGlyValIleGlnValLeuPheThrAlaValAlaLeuIleMet 313
Db 940 GATGAGCTATAGTTGGTGTGTGAGGCTCCGTGTGTGCTGATGCGCTGTCAACATG 999
QY 314 AspArgAlaGlyArgArgLeuLeuValLeuSerGlyValValMetValPheSerThr 333
Db 1000 GACCTGGCGGCGGAAAGTCTCTGTGTGTGTCAGCATCCATTATGTTTCTCCCAAC 1059
QY 334 SerAlaPheGlyAlaTyrPheLysLeuThrGlnGly-----GlyProGlyAsnSerSer 351
Db 1060 CTGACGCTGGGCTGTATGTCAGCTTGTGCCAAGGACTCTGACCCCAACAGACTGTG 1119
QY 352 HisValAlaIleSerAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAla 371
Db 1120 GAGATCGTGACGCTTGGGGGCGACGAGCAGCAGCCCGCGCGCTTCAACTATCTCAC 1179
QY 372 TrpLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyPro 391
Db 1180 CTGATACCCCTGCTGGCCACCAGTCTTTCATTATGGGCTATGCCATGGCTGGGGGCC 1239
QY 392 IleProTrpLeuLeuMetSerGluIlePheProLeuHisValLysGlyValAlaThrGly 411
Db 1240 ATCACTGGCTCTCATGTCTGAGGTTCGCCCTCGCGCGGTGGTGTGGCTCAGGG 1299
QY 412 IleCysValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeu 431
Db 1300 CTCTGGTGTCTGTGCTGAGCTGAGCTGAGCTGAGTCTGCTGAGTCTGCTGAGTCTG 1358
QY 432 MetGluValLeuArgProTyrGlyAlaPheTrpLeuAla-SerAlaPheCysIlePheSe 451
Db 1359 AGTGAATGCTTCGGCTCCAGGTGCTTCTTCTTCTGCGGCATCTGCTGCTGCTAG 1418
QY 451 rValLeuPheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleTh 471
Db 1419 CTGTCTCTTCAGAGCTGTGTGTGCTGAGCAGGCGCGCTGCTGAGCAGATCGA 1478
QY 471 rAlaHisPheGluGlyArg 477
Db 1479 GGCCTTCTTCCACACTCGC 1497

RESULT 12
ADA84076
ID ADA84076 standard; DNA; 2356 BP.
XX AC ADA84076;
XX DT 20-NOV-2003 (first entry)
XX DE Human SLC2A6 gene.
XX KW human; marker; expressed sequence tag; EST; arabidopsis; tumour;
XX stress-induced phenotype; hyperosmotic stress; colon cancer; immunogen;
XX vaccine; ds; gene.
XX OS Homo sapiens.
XX PN WO2002103028-A2.
XX PD 27-DEC-2002.
XX PF 30-MAY-2002; 2002WO-1B004189.
XX PR 30-MAY-2001; 2001US-0293999P.
```

```
PR 22-OCT-2001; 2001US-0330457P.
PR 19-FEB-2002; 2002US-0357144P.
XX (BIOM-) BIOMEDICAL CENT.
XX Baranova AV, Yankovsky NK, Kozlov AP, Lobashev AV, Krukovskaya IL;
XX WPI; 2003-175241/17.
XX P-PSDB; ADA84077.
XX Determining if a nucleic acid is a marker for a phenotype/cell type of
XX interest, by global comparison of expressed sequence tags known to be
XX expressed in the phenotype/cell type with all ESTs expressed in normal
XX tissue.
XX Claim 23; Page 457-458; 516pp; English.
XX The invention relates to a novel method for determining if a nucleic acid
XX is a marker for a predetermined phenotype/cell type of interest from a
XX biological species. The method comprises performing a global comparison
XX of a group of expressed sequence tags (ESTs) known to be expressed in the
XX phenotype/cell type of interest with all ESTs expressed in normal tissue
XX in order to identify ESTs that are preferentially expressed in the
XX phenotype/cell of interest. A method of the invention is useful for
XX determining whether a nucleic acid is a marker for a predetermined
XX phenotype or cell type of interest from a biological species, preferably
XX Arabidopsis or human. The cell type of interest is an abnormal cell such
XX as a tumour cell, and the predetermined phenotype is a stress-induced
XX phenotype such as hyperosmotic stress or high salt conditions. A method
XX of the invention is also useful for determining the progression of colon
XX cancer in a human, for detecting a tumour cell, and for regulating or
XX preventing the growth of a tumour cell. An antibody of the invention is
XX useful for detecting the absence or presence of peptides encoded by
XX tumour-associated markers. A polypeptide of the invention is useful as an
XX immunogen for vaccinating an animal. The present sequence encodes a
XX tumour-associated antigen of the invention.
XX Sequence 2356 BP; 409 A; 786 C; 721 G; 440 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 1.18e-73 Length: 2356
Score: 958.50 Matches: 227
Percent Similarity: 59.22% Conservatives: 78
Best Local Similarity: 44.08% Mismatches: 164
Query Match: 39.01% Indels: 47
DB: 7 Gaps: 12
```

US-09-886-954A-1 (1-477) x ADA84076 (1-2356)

```
QY 3 ProGlu---AspProGluGluThrGlnProLeuLeuGly----- 14
Db 12 CCCGAGAGAGACCCGGCCATGCAGGAGCGCTGTGGAGCGCGAGGCCGACTACGAC 71
QY 15 -----ProProGlyGlySerAlaProArgGlyArg----- 24
Db 72 ACCTTCCCGAGAGAGCGCCCGCTCGCAGGGGACAGGGCGCGTCTGGACCTCGAC 131
QY 25 -----ArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPhe 42
Db 132 AACAAAGGGTGTCTCTGGCCACCTTCGCCGAGTCTCGGCAATTTCAGCTTTGGGTAT 191
QY 43 AlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaPro 62
Db 192 GCCTTGGTCTACACATCCCTGTCATCCAGCCCTGGAGCGCTCTTGGATCTGACCTG 251
QY 63 ArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAla 82
Db 252 CATCTGACCAATCCAGGCATCTCTGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 311
QY 83 GlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeu 102
Db 312 GGAGGCTGAGTGCCATGATCTCTCAACGACCTCTCTGGGCGGAGAGCTGAGCATCATG 371
```



```
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 38..1561
FT     /*tag= a
FT     /product= "Human SLC2A"
XX
XX WO200298467-A1.
XX
XX PD 12-DEC-2002.
XX
XX 03-JUN-2002; 2002WO-US017419.
XX
XX 05-JUN-2001; 2001US-0296076P.
XX
XX 10-OCT-2001; 2001US-0328605P.
XX
XX 15-FEB-2002; 2002US-0357253P.
XX
XX (EXBL-) EXELIXIS INC.
XX
XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX
XX WPI; 2003-201283/19.
XX
XX P-PSDB; ABP58363.
XX
XX Identifying candidate p53 pathway modulating agent as therapeutic target
XX for disorders of defective p53 function e.g. cancer, by assaying purified
XX solute carrier family 2 (SLC2A) polypeptide or nucleic acid with a test
XX agent.
XX
XX PS Disclosure; Page 38-39; 58pp; English.
XX
XX The present sequence is that of human solute carrier type 2A (SLC2A)
XX nucleic acids 12735203 and 9938031. In the present invention, genetic
XX screens were designed to identify modifiers of the p53 pathway in
XX Drosophila in which p53 was overexpressed. Human orthologues
XX (polynucleotides and polypeptides) of one such modifier were then
XX identified, including nucleic acids 12735203 and 9938031. SLC2As are
XX glucose transporter proteins with sugar transporter domains. SLC2A
XX nucleic acids and polypeptides are attractive drug targets for the
XX treatment of pathologies associated with a defective p53 signalling
XX pathway, such as cancer. The invention provides in vitro and in vivo
XX methods of assessing SLC2A function. Modulation of an SLC2A or its
XX binding partners is useful for understanding the association of the p53
XX pathway and its members in normal and disease conditions and for
XX developing diagnostic and therapeutic modalities of p53-related
XX pathologies. SLC2A-modulating agents that act by inhibiting or enhancing
XX SLC2A expression, directly or indirectly, e.g. by affecting an SLC2A
XX function such as transport or binding activity, can be identified using
XX methods provided. Modulators include small molecules, nucleic acids,
XX antibodies, antisense oligonucleotides and phosphothioate morpholino
XX oligomers (claimed)
XX
XX SQ Sequence 2487 BP; 443 A; 826 C; 752 G; 466 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1-27e-73 Length: 2487
XX Score: 958.50 Matches: 227
XX Percent Similarity: 59.22% Conservatives: 78
XX Best Local Similarity: 44.08% Mismatches: 164
XX Query Match: 39.01% Indels: 47
XX DB: 7 Gaps: 12
XX
XX US-09-886-954A-1 (1-477) x ABZ24790 (1-2487)
XX
XX QY 3 ProGlu---AspProGluGluThrGlnProLeuLeuGly----- 14
XX
XX DB 20 CCGAGAGAGACCCGCCCATGCGAGGCGCGTCTGGGAGCGCGAGGCCGCCGACTACGAC 79
XX
XX QY 15 -----ProProGlyGlySerAlaProArgGlyArg----- 24
XX
XX DB 80 ACCTTCCCGAGAGAGCCGCCCGCGTCCGAGGGGACAGGGCGCGGTCGGACCCCTCGAG 139
XX
XX QY 25 -----ArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPhe 42
```

```
Db 140 AACAAAGGGTTCCTGSCACCTTGGCCGAGTCTCGCAATTTCAGCTTTGGGTAT 199
QY 43 AlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaPro 62
Db 200 GCCCTGGTCTACACATCCCTGTTCATCCAGCCCTGGAGCGCTCTTGGATCCTGACCTG 259
QY 63 ArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValTrpLeuGlyAlaAlaAla 82
Db 260 CATCTGACCAATCCAGGCATCTCTGGTTGGTCCGTTGTTCACTCCCTGGAGCAGCGGCC 319
QY 83 GlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeu 102
Db 320 GGAGGCTGAGTGCCATGATCTCAACGACCTCTGGGCCGGAAGCTGAGCATCATGTTTC 379
QY 103 CysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
Db 380 TCAGCTGTGCGTGGCGGCGGCTATCGCTCATGGCGGGTGGCCAGCGGCTCTGGATG 439
QY 123 LeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaPro 142
Db 440 CTGCTGCTCGAAGAGACGCTGACGGGCTTCGCCGGGGGCTCACAGCTGCCTGCATCCG 499
QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
Db 500 GTGTAGCGTGTCTGAGATTGCTCCGCCAGCGTTCGTGGGCTCTGGGGGCCACACCCAG 559
QY 163 LeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArg 182
Db 560 CTCATGGCAGTGTTCGGATCTCCTGCTCTACGCCCTTGGCCTCTGCTGCGGTGGCG 619
QY 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
Db 620 TGGCTGGTGTGGCCGGAGCGCTGTGCTCATCATGATCTCTGCTCAGCTTCATG 679
QY 203 ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaLeu 222
Db 680 CCCAACTCGCGCGCTTCTGCTCTCTCGGGCAGGACGAAAGAGCGCTCGGGCGCTG 739
QY 223 ArgPheLeuTrpGlySerGluGlnGly-----TrpGluAspProProIleGly----- 238
Db 740 GCCTGGCTGGTGGGACGAGTGCATGTCACCTGGAGTTCGACAGATCAGACAAAC 799
QY 239 ---AlaGluGlnSerPheHisLeu-----AlaLeuLeuArgGlnProGlyIleTyrLys 255
Db 800 GTCCGGAGACAGACAGCGCGAGTATCGTGGGCTGAGGACGCGGCCCCACACGTGCGCG 859
QY 256 ProPheIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaVal 275
Db 860 CCCATCACGCTGGCGCTGTGATGCGCTCTCTGACAGAGTGCAGGGCATCACGCCCATC 919
QY 276 MetPheTyrAlaGluThrIlePheGluGluAlaLysPhe-----LysAspSer 291
Db 920 CTGGTCTACCTGCAGTCCATCTTCGACGACACCGTGTCTGCTGCCCCCAGGACGAC 979
QY 292 SerLeuAlaSerValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeu 311
Db 980 -----GCAGCCATCGTTGGGCGGTGGCGGCTCTGTCGCTGTGATCGCGCCCTC 1030
QY 312 IleMetAspArgAlaGlyArgArgLeuLeuLeuValLeuSerGlyValValMetValPhe 331
Db 1031 ACCATGGACCTCGAGGCGCGAAGGTGTGCTCTCTGCTCTCAGCGGCCATCATGTTGCT 1090
QY 332 SerThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSer 351
Db 1091 GCCAACCTGACTCTGGGCTGTACATCCACTTT-----GGCCCCAGGCGCTCTGAGC 1141
QY 352 HisValAlaIleSerAlaProValSer-----AlaGlnProValAsp 365
Db 1142 CCCAAGACACTCGCGGCCCTGGAAGCGAGTCTGCGGGGAGCTTGGCGCAGCGCCCTGGA 1201
QY 366 AlaSerValGly---LeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAlaGly 384
```

```
Db 1202 GCACCCGCTGCTACCTCACCCCTGGTGCCCTGCTGGCCACCATGCTCTTTCATCATGGGC 1261
Qy 385 PheAlaValGlyTrpGlyProIleProTrpLeuMetSerGluIlePheProLeuHis 404
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
1262 TACGCCGTGGGCTGGGGTCCCATACCTGGGTGCTCATGTCTGAGGTCTGCCCTGGCTGGT 1321
Qy 405 VallysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeuVal 424
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
1322 GCCCGTGGCGTGGCTCAGGGCTCTGGGTGCTGGCCAGCTGGCTACCGCCTTGGTCCTC 1381
Qy 425 ThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeuAla 444
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
1382 ACCAAGTCCTTCCTGCCAGTGGT-GAGCACCTTCGGCCTCCAGGTGCCCTTCTTCTTCTT 1440
Qy 445 -SerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrLysGl 464
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
1441 CGCGGCCATCTGCTGTGGTGGCTGGTGTTCACAGGTGCTGTGTGCCCGAGACCAAGGG 1500
Qy 464 yLysThrLeuGluGlnIleThrAlaHisPheGlu---GlyArg 477
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
1501 ACGGTCCCTGGAGCAGATCGAGTCCTTCTCCGCATGGGGAGA 1543
```

Search completed: September 28, 2004, 02:12:01
Job time : 518 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 28, 2004, 01:55:59 ; Search time 111 Seconds
(without alignments)
2384.789 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPEDPBTQPLLPGPGSA.....CVPETKGLQIHTAFEGR 477

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09886954/runat_27092004_170632_13412/app_query.fasta_1.647
-DB-Issued Patents NA -QMT-fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=-1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09886954.cgn2_1_105@runat_27092004_170632_13412 -NCPUS=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	605	24.6	1545	4	US-09-489-039A-4731
2	578.5	23.5	1431	4	Sequence 4731, Ap
3	570.5	23.2	3915	4	Sequence 4762, Ap
4	542.5	22.1	1752	4	Sequence 1104, Ap
5	532.5	21.7	1675	4	Sequence 17, Appl
6	518	21.1	1960	4	Sequence 15, Appl
7	514.5	20.9	2856	4	Sequence 135, App
8	514.5	20.9	2856	4	Sequence 135, App
9	514.5	20.9	2856	4	Sequence 135, App
10	514.5	20.9	2856	4	Sequence 135, App
11	514.5	20.9	2856	4	Sequence 135, App
12	507.5	20.7	1506	4	Sequence 4560, Ap

13	506.5	20.6	1914	4	US-09-291-922-19	Sequence 19, Appl
14	499	20.3	2592	4	US-09-591-025-8	Sequence 8, Appl
15	498	20.3	5227	4	US-09-919-172-79	Sequence 79, Appl
16	495	20.1	2592	4	US-09-894-927B-8	Sequence 8, Appl
17	489	19.9	1566	4	US-09-489-039A-2378	Sequence 2378, Ap
18	481	19.6	272	4	US-09-016-434-855	Sequence 855, App
19	478	19.5	1689	4	US-10-162-012-45	Sequence 45, Appl
20	478	19.5	4385	4	US-10-162-012-43	Sequence 43, Appl
21	468.5	19.1	2089	4	US-09-291-922-25	Sequence 25, Appl
22	468.5	19.1	3000	2	US-08-928-692-9	Sequence 9, Appl
23	468.5	19.1	3000	4	US-09-339-972-9	Sequence 9, Appl
24	464.5	18.9	1853	4	US-09-291-922-23	Sequence 23, Appl
25	458.5	18.7	1872	4	US-09-291-922-27	Sequence 27, Appl
26	453	18.4	4411529	3	US-09-103-840A-1	Sequence 1, Appl
27	446	18.2	2017	4	US-09-291-922-21	Sequence 21, Appl
28	439.5	17.9	1776	4	US-09-679-686B-11	Sequence 11, Appl
29	429	17.5	4403765	3	US-09-103-840A-2	Sequence 2, Appl
30	428	17.4	272	4	US-09-016-434-440	Sequence 440, App
31	423	17.2	972	4	US-09-134-001C-1798	Sequence 1798, Ap
32	398.5	16.2	2343	2	US-09-031-392-1	Sequence 1, Appl
33	398.5	16.2	2343	3	US-09-299-549-1	Sequence 1, Appl
34	398.5	16.2	2343	4	US-09-610-417-1	Sequence 1, Appl
35	382.5	15.6	3035	4	US-09-833-381-1987	Sequence 1987, Ap
36	368.5	15.0	2842	4	US-09-620-312D-574	Sequence 574, App
37	339	13.8	2601	4	US-09-291-922-7	Sequence 7, Appl
38	337	13.7	918	4	US-09-107-532A-561	Sequence 561, App
39	296.5	12.1	1487	4	US-09-291-922-13	Sequence 13, Appl
40	293.5	11.9	2824	4	US-09-291-922-1	Sequence 1, Appl
41	289.5	11.8	1009	4	US-09-291-922-15	Sequence 15, Appl
42	281	11.4	1692	4	US-09-291-922-9	Sequence 9, Appl
43	276	11.2	1896	3	US-08-501-572-6	Sequence 6, Appl
44	276	11.2	1896	3	US-09-040-444-6	Sequence 6, Appl
45	257.5	10.5	1341	4	US-09-543-681A-2407	Sequence 2407, Ap

ALIGNMENTS

RESULT 1
US-09-489-039A-4731
; Sequence 4731, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4731
; LENGTH: 1545
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4731

Alignment Scores:
Pred. No.: 5.87e-53 Length: 1545
Score: 605.00 Matches: 156
Percent Similarity: 51.60% Conservative: 86
Best Local Similarity: 33.26% Mismatches: 187
Query Match: 24.62% Indels: 40
DB: 4 Gaps: 9

US-09-886-954A-1 (1-477) x US-09-489-039A-4731 (1-1545)

Qy 20 AlaProArgGlyArgArg-----ValPheLeuAlaAlaPheAlaAla 34
Db 154 TCGCGCGGAGCGCAACGTGATACCCGCGGATGAACCTGGTTTGTTCATCGCGCGCG 213
Qy 35 LeuGlyProLeuSerPheGlyPheAlaLeuGlyTyrSerProAlaIleProSerLeu 54

Db 214 GTGGCTGGCTTCTTTGGCTGGATATCGCGGTGATATCGGAGCGTTGCCCTTTATA 273

Qy 55 GlnArgAlaAlaProProAlaProArgLeuAspAlaAlaSerTrpPheGlyAla 74

Db 274 ACCGACCATTTTACCTTATCCAGCCAGCTTTCAGGAG-----TGGGTGGTTAGC 321

Qy 75 ValValThrLeuGlyAlaAlaGlyGlyGlyValLeuGlyGlyTrpLeuValAspArgAla 94

Db 322 AGCATGATGTTGGGGGGGGGATAGCGCGCTGTTTAAAGCGTGGCTGCTCTTTCGCGCTT 381

Qy 95 GlyArgLysLeuSerLeuLeuLeuCysSerValProPheValAlaGlyPheAlaValIle 114

Db 382 GGGCGTAAATACAGCTGATGCGGGGGCGGTACTCTTTTGTGGCGCGCTCTATCGATCC 441

Qy 115 ThrAlaAlaGlnAspValTrpMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCys 134

Db 442 GCTTTTGGCGCAGCGTGGAGTGTCTGTGTGGCGCGCTGTGTGGCGGTGGCAGTC 501

Qy 135 GlyValAlaSerLeuValAlaProValTrpIleSerGluIleAlaAlaTrpProAlaValArg 154

Db 502 GGGATGCTCTTATACCGCGCGCTGTACCTCTCGAGATGCCAGCAACGTGGCG 561

Qy 155 GlyLeuLeuGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAlaTrpLeu 174

Db 562 GGGAAATGATCAGTATGATCAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 621

Qy 175 AlaGlyTrpValLeuGluTrp-----ArgTrpLeuAlaValLeuGly-----CysVal 190

Db 622 TCCGATACCGCTTTAGCTACAGCGGTAACTGGCGCGCATGCTGGCGGTGGCGGTG 681

Qy 191 ProProSerLeuMetLeuLeuMetCysPheMetProGluTrpProArgPheLeuLeu 210

Db 682 CGGCGGTGATCTCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741

Qy 211 ThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuTrpGlySerGluGln 230

Db 742 GAGAAAGGAGCCATATCGAAGCGGAGAGTGTGCGGATGCTGCGGATACCTTCGGAA 801

Qy 231 GlyTrpGluAspProIleGlyAlaGluGlnSerPheHisLeu----- 245

Db 802 AAGGCGCGAGCTTAAACGAGATCGGTGAGAGCCCTCAAGCTGAAGCAGCGCGGTGG 861

Qy 246 AlaLeuLeuArg---GlnProGlyIleTrpPheIleIleGlyValSerLeuMet 264

Db 862 GGTGTTGTTAAGATCAATCGTAACGTGCGCGCGGTGCTTCTGCGCATGCTGCGAG 921

Qy 265 AlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIlePheGlu 284

Db 922 GCGATGCGACCAATTCACCGCATGAACATCATCATGCTACTATGCGCGCGGTATCTTAA 981

Qy 285 GluAlaLysPhe-----LysAspSerSerLeuAlaSerValValGlyValIle 301

Db 982 ATGGCGCGCTTTACACTACTGAACAGAGATGATCGCCACCTGGTGGTGGCGCTGACC 1041

Qy 302 GlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeu 321

Db 1042 TTTATGTTTGGACCTTTATTCGGTGTTCACAGTGATAAAGCGGCGCGCAAGCCAGC 1101

Qy 322 LeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTrpPheLys 341

Db 1102 CTGAAATTCGGCTTTAGCGTGATGCGCTGCGCACTCTGGTGGCGG-----TACTGC 1155

Qy 342 LeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAla 361

Db 1156 CTGATGCAATTCGAAATGGAC----- 1179

Qy 362 GlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPhe 381

Db 1180 -----GCATCAGCGCGCTCTCTCGTGGCTCTCGTGGCATGACCATGATGCT 1227

Qy 382 IleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePhe 401

Db 1228 ATTGCGGGGTATGCGATGCGCGCGCGGTGGTGGATCCTCTGCTCGAGATCCAG 1287

Qy 402 ProLeuHisValLysGlyValAlaAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAla 421

Db 1288 CCGCTAAATGCGGAGCTTCGGTATCACTGCTCGACACACCACTGGGTGTCGAC 1347

Qy 422 PheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTrpGlyAlaPhe 441

Db 1348 ATGATCATCGCGCCACTTTCTGACGCTGCTTGACGCGATTGGCGCGCGGACCTTC 1407

Qy 442 TrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGlu 461

Db 1408 TGGCTCTACACGGCGCTCAAGCTGGCTTTATCGGCATCAGCTTCTGCTGATCCCGAA 1467

Qy 462 ThrLysGlyLysThrLeuGluGlnIle 470

Db 1468 ACCAAGATGTACCCCTCGACACATT 1494

RESULT 2

US-09-489-039A-4762

; Sequence 4762, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709,2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 4762

; LENGTH: 1431

; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-4762

Alignment Scores:

Pred. No.: 2,92e-50 Length: 1431

Score: 578.50 Matches: 149

Percent Similarity: 48.74% Conservative: 84

Best Local Similarity: 31.17% Mismatches: 180

Query Match: 23.54% Indels: 65

DB: 4 Gaps: 10

US-09-886-954A-1 (1-477) x US-09-489-039A-4762 (1-1431)

Qy 27 PheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGlyTyr 46

Db 82 TTCCTTCGCTCTTCTTCCTCGCGCGCTGGCTGGCTGCTGCTGGCTTGGATATCGGTGT 141

Qy 47 SerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArgLeuAsp 66

Db 142 ATTCGGGTGCTTACCTTATTCCTCAATCAGTTC-----CAGATTTCCGCC 189

Qy 67 AlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAlaGlyValLeu 86

Db 190 CACACCGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 249

Qy 87 GlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuLeuLeuLeuLeu 106

Db 250 AGCGCTGGCTCTCTTCAAACTGGCGCGGAAAAAGAGCTGATGATCGCGCGCATCCTC 309

Qy 107 PheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGly 126

Db 310 TTCGTGCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 369

Qy 127 ArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaProValTrpIleSer 146

Db 370 CGTGTGCTGCTCGGCTGGCGGTGGCGCTCGCTCATATACGGCTCCGCTGATCTGTGCG 429

Qy 147 GluIleAlaTrpProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetVal 166

```

Db 430 GAAATGCGCCCAAGAAAAATTCGCGCAGTATGATTTCCATGATACCAGCGTGTATGATCACC 489
Qy ValGlyIleLeuLeuAlaValGlyTrpValLeuGlu-----TrpArg 182
Db 490 ATCGGATCTTGGCGCTATCTCTGTACACCGTTTCAGTACAGCGCGCATGGCGC 549
Qy TrpLeuAlaValLeuGlyCysValProSerLeuMetLeuLeuMetCysPheMet 202
Db 550 TGGATGCTCGGGGTTATCATCATTCGCGGGTTTGTCTGTATCGCGGTATCTTCG 609
Qy ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaLeu 222
Db 610 CCGGACAGCCCGCTGTGTTCCGCCCAACGTCGCTTGTGATGCGGACGCGTGTG 669
Qy 223 -----ArgPhe 224
Db 670 CTGCGCTGCGGATACAGCGCCGGAAGCGAGCTCGATGAATCCGTGAAAGC 729
Qy 225 LeuTrpGlySerGluGlnGlyTrpGluAspProPheGlyAlaGluGlnSerPheHis 244
Db 730 CTGAAGGTAAACAGTCCGGCTG-----753
Qy 245 LeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIle---IleGlyValSerLeu 263
Db 754 ---TCGCTGTTTAAAGACAAACAGCAACTTCGCGCGCGGTTCCTCGCATCTGCTG 810
Qy 264 MetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIlePhe 283
Db 811 CAGGTGATGCAACAGTTCACCGGATGAACGTATCATGTACTACGCGCCGGAAGATCTTT 870
Qy 284 GluGluAlaLysPheLysAspSer-----LeuAlaSerValValGlyVal 300
Db 871 GAGCTGCGGGTTATGCCAACACCACTGAGCAATGTGGGGACAGTATGCTGGTCTG 930
Qy 301 IleGlnValLeuPheThrAlaValAlaLeuLeuMetAspArgAlaGlyArgArgLeu 320
Db 931 ACTAACGTGTGGCCACTTTATGCCATCGCTGTGTGTCGACCGTGGGCGGTAAACCG 990
Qy 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
Db 991 ACGTGTATCTTGGCTTATGCTGATG-----GCCGCGGA-----1026
Qy 341 LysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
Db 1027 -----ATGGCGCTCTGGTACCATGATGACATCGCATCTCTCTACCGCC 1077
Qy 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
Db 1078 -----CAGTACATCGCGCTCTGATGCTGCTGATG 1107
Qy 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400
Db 1108 TTATCTGCTGGTTTCGCATGAGCGCGCCCACTGATTTGGGTACTGTGCTCGAATC 1167
Qy 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
Db 1168 CAGCGCGTGAAGCGCGCATCTCGGTATCACCTGCTCCACAGCGACCACTGGATTGCC 1227
Qy 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
Db 1228 AACATGATTGCGCGCCACTCTCTGACCATGCTCAACTCGCTGGGCGAGCGCAATACC 1287
Qy 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
Db 1288 TTCTGGGTGACGGCGTCTGAACGTGCTGTTTATCTCTGTGACGTGTGCTGATCCCG 1347
Qy 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHis---PheGluGlyArg 477
Db 1348 GAAACCAAAAACGCTCTCGCTGGACATATTGAACGTAACTGATGAGGGTCTGT 1401

```

RESULT 3

US-09-023-655-1104

; Sequence 1104, Application US/09023655

```

; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1104:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3915 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g183684
; US-09-023-655-1104

```

Alignment Scores:

Pred. No.:	9,586-49	Length:	3915
Score:	570.50	Matches:	152
Percent Similarity:	49.90%	Conservative:	88
Best Local Similarity:	31.60%	Mismatches:	190
Query Match:	23.22%	Indels:	51
DB:	4	Gaps:	9

US-09-886-954A-1 (1-477) x US-09-023-655-1104 (1-3915)

Qy	26	ValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGly	45
Db	270	CTGATATTTTCCATCAGTTGCTACATCGGCTCTTCCAAATTGGCTACACACTGGG	329
Qy	46	TyrSerSerProAlaIleProSerLeuGlnArg-----	56
Db	330	GTCCATCAATGCTCTCTGAGAAGATCATAAAGGAATTTATCAATAAACTTTTGACGGCAAG	369
Qy	57	---AlaAlaProAlaProArgLeuAspAlaAlaSerTrpPheGlyAlaVal	75
Db	390	GGAATGCCCCACCCTCTGAGTGTGCTACGCTCTCTGTGGTCTGTGTGTGGCCATA	449
Qy	76	ValThrLeuGlyAlaAlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGly	95
Db	450	TTTTCGTCGGGGTATGATCGGCTCTTTCCGTCGGACTCTTCGTCACACGCTTTGGC	509
Qy	96	ArgLysLeuSerLeuLeuLeuCysSerValProPheValAlaGly-----PheAla	112

Db 510 AGGCGAAATCAATGCTGATGTAACCTGTTGGCTGCTCACTGGTGGCTTTATGGGA 569
 Qy 113 ValIleThrAlaAlaGlnAspValTrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeu 132
 Db 570 CTGTGTAAGTAGCTAAGTCGGTGAATGCTGATCTCGGGTCGCTGTTGTTATGGCTC 629
 Qy 133 AlaCysGlyValAlaSerLeuValAlaProValTyrIleSerGluIleAlaValProAla 152
 Db 630 TTCTGGCGACTGACAGGTTTGGCCCATGTATCTGGAGAGATCTCGCTACTGCC 689
 Qy 153 ValArgGlyLeuLeuGlySerCysValGlnLeuMetValValGlyIleLeuLeuAla 172
 Db 690 CTGCGGGTGCCTTTGGCACTCTCAACAGCTGGGCATCGTTGTTGGAATCTGCTGCC 749
 Qy 173 TyrLeuAlaGly-----TrpValLeu-----GluTrpArgTrpLeuAlaValLeuGly 188
 Db 750 CAGATCTTTGGTCTGGAATTCATCTTGGTGTGAAGAGCTATGGCCGCTGCTACTGGGT 809
 Qy 189 -----CysValProSerLeuMetLeuLeuMetCysPheMetProGluThrPro 206
 Db 810 TTTACCATCTTCTGCTATCTCAACAGTGCAGCCCTCCATTTTGGCTGAAAGTCCC 869
 Qy 207 ArgPheLeuLeuThrGlnHisArgArgGlnGlu---AlaMetAlaAlaLeuArgPheLeu 225
 Db 870 AGATTTTGTCTATTACAGAAAAGAGAGAGATGCTTAAGCAGATCTCTCCAGCGGTG 929
 Qy 226 TrpGlySerGlnGlnTrpGluAspProProlIleGlyAlaGluGlnSerPheHis--- 244
 Db 930 TGGGGACCCAGGATGTPATCCCAAGACATCCAGAGATGAAAGATGAGAGTCAAGGATG 989
 Qy 245 -----LeuAlaLeuLeuArgGlnProGlyIleTyrLysPro 256
 Db 990 TCACAGAAAAGCAAGTCACGGTCTAGAGCTTTTAGAGTGTCAGCTACCGACAGCC 1049
 Qy 257 PheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMet 276
 Db 1050 ATCATCATTTCCATTGTGCTCCAGCTCTCTCAGCAGCTCTCTGGGATCAATGCTGTTC 1109
 Qy 277 PheTyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerVal 296
 Db 1110 TATTACTCAACAGGAATCTTCAAGAGTGCAGGTGTTCAAGAGCCCATCTATGCCACCAT 1169
 Qy 297 ValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAla 316
 Db 1170 GGCGGGTGTGGTTAATACTATCTTCACTGTAGTTCTTCTTCTGCTGGAAAGGCA 1229
 Qy 317 GlyArgArgLeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPhe 336
 Db 1230 GGAAGAAGGACTCTGCATATGATA----- 1253
 Qy 337 GlyAlaTyrPheLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSer 356
 Db 1254 -----GGCCTTGGAGGATGGCTTTTGTTCACGCTC 1286
 Qy 357 AlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGly 376
 Db 1287 ATGACTGTTCTTTGTTATTAAGGATAACTATAATGGGATGAGCTTGTCTGTTATGG 1346
 Qy 377 SerMetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeu 396
 Db 1347 GCTATCTTGGTCTTTGTAGCTCTTTGAAATGGACAGGCCCATCTCCCTGGTTTATT 1406
 Qy 397 MetSerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThr 416
 Db 1407 GTGGCCGAACACTTTTCAGCCAGGGCCCCCGCCAGCTGCGATGCGAGTGGCCGGCTGCTCC 1466
 Qy 417 AsnTrpLeuMetAlaPheLeuValThrLysGlnPheSerSerLeuMetGluValLeuArg 436
 Db 1467 AACTGGACCTCAACTTCTTAGTGGATGCTGCTTCCCTCCCTGCTCACTATTATTAGGA 1526
 Qy 437 ProTyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeu 456

```
QY 138 SerLeuValAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeu 157
Db 516 AATCATGCTGTGGCGGTGTACCTGTGGAGATGGCGCTCTCCGGGGCATGCTC 575
QY 158 GlySerCysValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeu----- 174
Db 576 AACATCGGTTCCAGCTCATGATCACCATCGGCATCTGTGGCGGGCTCATCAATTAC 635
QY 175 -----AlaGlyTyrValLeuGluTyrArgTyrPheLeuAlaValLeuGly 188
Db 636 GACACCAACAAGATCAAGGCGGGTAC-----GGTGGCGCATACGCTGGCCATCGCG 699
QY 189 CysValProProSerLeuMetLeuLeuMetCysPheMetProGluThrProArgPhe 208
Db 690 GCGTCGCGGCGGCGCATCATCACCCTGGGTCTTTTCTCCCGACACCCCACTCC 749
QY 209 LeuLeuThrGlnHisArgGlnGluAlaMetAlaAlaLeuArgPheLeuTyrPhe 228
Db 750 CTATCGAGGTGGCCACCGGAGGCGCGCGCGCATGCTCAACCGCATCGCGGCAGC 809
QY 229 Glu-----GlnGlyTyrPheGluAspProProIleGlyAlaGluGlnSerPheHis 244
Db 810 GACGTGGACATCAGCGAGGATACGCGGACCTGGTGGCGAGCGAGGAGTCC----- 863
QY 245 LeuAlaLeuLeuArgGlnPro-----GlyIleTyrLysPro---Phe 257
Db 864 ---AAGCTGTGTGACGACCGCGTGGCGCAATCTTCGCGCAAGTACCGGCGCCAGCTG 920
QY 258 IleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPhe 277
Db 921 ACCATGGCGATCATGATCCCTTCTTCAGCAGCTGACGGGCATCAACGTCATCATGTT 980
QY 278 TyrAlaGluThrIlePheGluGluAlaLysPheLys---AspSerSerLeuAlaSer--- 295
Db 981 TAGCGCGCGTCTGTTCGAGAGCGTGGGTTCAGGGCGCGCGCGTCTCATGTCGCGCC 1040
QY 296 ValValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArg 315
Db 1041 GTCATCGGGCGCTGTCAACGTGTTCGCGACGCTCGTTCGCTGTTCACCGTCCACCGG 1100
QY 316 AlaGlyArgArgLeuLeuValLeuSerGlyValValMetValPheSerThrSerAla 335
Db 1101 CTGGTCTCCCGAAGCTGTTCGTGCGGCGCGCGACGACATGCTCTGAGCGCAGTGTG 1160
QY 336 PheGlyAlaTyrPheLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIle 355
Db 1161 GTGGGCACTGTGCGGTCAAGTTCGGGACGCGGC----- 1199
QY 356 SerAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTyrPheAlaVal 375
Db 1200 -----GTGGGGGAGATGCCC-----AAGGGGTACGCGCGCGCGGTGGTG 1238
QY 376 GlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTyrPheProIleProTyrLeu 395
Db 1239 CTCTTCATCTGCTCTATGTCGGGGTTCGCTGGTGTGGTGGTGGTGGTGGTGGTGGT 1298
QY 396 LeuMetSerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeu 415
Db 1299 GTGCCACGAGATCTTCGCGTGGAGATCAGCGCGCGGGGCGAGACATCAACGTGTG 1358
QY 416 ThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeu 435
Db 1359 GTGAACATGTCTTTCATCTTCGTATCGCGCGCGGTTCCTCACCATGCTCTGCCACATG 1418
QY 436 ArgProTyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThr 455
Db 1419 AAG---TTCGGCTCTTCTACTTCTTCGCGGTGGTGGTGGTGGTGGTGGTGGTGGT 1475
QY 456 LeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGlu 475
Db 1476 GCGCTCTCTTCGCGGAGACCAAGACGTGCCCATCGAGGAGATGGTGTCTGTGGAAG 1535
QY 476 Gly 476
```

```
Db 1536 GGA 1538
RESULT 5
US-09-679-686B-1
; Sequence 1: Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
; FILE REFERENCE: BB1160 US NA
; CURRENT APPLICATION NUMBER: US/09/679,686B
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/081,131
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: PCT/US99/07561
; PRIOR FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1675
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1624)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1638)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1670)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1674)
; OTHER INFORMATION: n = a, c, g or t
; US-09-679-686B-1
```

```
Alignment Scores:
Pred. No.: 2,21e-45 Length: 1675
Score: 532.50 Matches: 147
Percent Similarity: 44.89% Conservative: 90
Best Local Similarity: 27.84% Mismatches: 190
Query Match: 21.67% Indels: 101
DB: 4 Gaps: 13
```

US-09-886-954A-1 (1-477) x US-09-679-686B-1 (1-1675)

```
QY 7 GluGluThrGlnProLeuLeuGlyProGlyGlySerAlaProArgArgArg--- 25
Db 150 GAGCAGCGAGATGCGCGCGGGTGTGGCGGGCGCGCGCGCGCGCGCGCGCG 209
QY 26 -----ValPheLeuAlaAlaPheAlaAlaLeuGly 36
Db 210 CTCTACGAGGCAAGATCACCGCTACTTCATCTCGCTTCATCGCTCGCTCGGC 269
QY 37 ProLeuSerPheGlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArg 56
Db 270 GGATCCCTCTTCGGCTATGACCTCGGAGTCTCAGC----- 305
QY 57 AlaAlaProAlaProArgLeuAspAlaAlaSerTrpPhe----- 72
Db 306 -----GGCGTGACTTCCATGGACGACTTCTCGTGAAGTCTTCCCGGAGCGTGAC 356
```

Qy 72 ----- 72
Db 357 CGCGGAAGCAGGCGACCTGCACGAGACGGACTACTGCAAGTACGACCAACACCGTGTG 416
Qy 73 -----GlyAlaValValThrLeuGlyAlaAla 82
Db 417 ACCTGTTTCACTCGTCTGCTTACTTCGCGGCTCTGCTCTCACTTCGCGGCTCC--- 473
Qy 83 GlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeu 102
Db 474 -----TACGTGACCAAGCGCCAGCGCGCGCGCGCATCATGGT 515
Qy 103 CysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
Db 516 GCGCGCGCGAGCTTCTTCGCGCGCGCCATCAACGCGCGCGCATGAACATGCCCATG 575
Qy 123 LeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaAspLeuValAlaPro 142
Db 576 CTCATCTGCGAGCATCTCTCGGCTGCGGCTCGGCTTCGCAATCAGGCGCGTGCCT 635
Qy 143 ValTrpIleSerGluIleAlaTrpProAlaValArgGlyLeuLeuGlySerCysValGln 162
Db 636 GTGTACCTGTGGAGATGGCGCGCGCTCTCCGCGGCATGCTCAACATCGCTTCAG 695
Qy 163 LeuMetValValGlyIleLeuLeuAlaTrpLeu----- 174
Db 696 CTGATGATCACCATCGGCATCTCGCGCGGAGCTCATCACTACGGCACCAACAGATC 755
Qy 175 ---AlaGlyTrpValLeuGluTrpArgTrpLeuAlaValLeuGlyCysValProSer 193
Db 756 AAGCGCGGTAC-----GGTGGCGGTGAGCTTGGCGCTGGCGGCTGGCGGCGC 809
Qy 194 LeuMetLeuLeuMetCysPheMetProGluThrProArgPheLeuLeuThrGlnHis 213
Db 810 ATCATCACCTGGCTCCCTCTCTCCCGGACACCCCACTCGCTGCTGGAGCGGCG 869
Qy 214 ArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuTrpGlySerGlu-----Gln 230
Db 870 CACCGGAGGAGCGACCGCGCATGCTCCGCGCGATCCGCGGCACGACATCGCGCAG 929
Qy 231 GlyTrpGluAspProIleGlyAlaGluGlnSerPheHisLeuAla----- 246
Db 930 GAGTACCGGACTGTGTGGCGCCAGCAGGAGGCGCGCGAGTGGCCACCGTGGCGG 989
Qy 247 ---LeuLeuArgGlnProGlyIleTrpLys---ProPheIleGlyValSerLeuMet 264
Db 990 AACATCTCGCGCGCG-----TACCGCGCGAGCTCACCATGGCGCGTCCGATCCCC 1043
Qy 265 AlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTrpAlaGluThrIlePheGlu 284
Db 1044 TTCTTCAGAGCTCAAGGGGATCAACGTATCATGTTTACGCGCGCGTGTGTTCGAC 1103
Qy 285 GluAlaLysPheLysAspSerSer-----LeuAlaSerValValGlyValIleGln 302
Db 1104 ACCTGGATTAAGAAAGAGCTTCTCTCCGTATCATCGGCGCTCGTCAAC 1163
Qy 303 ValLeuPheThrAlaValAlaAlaLeuMetAspArgAlaGlyArgArgLeuLeuLeu 322
Db 1164 GTCTTCGCCACCGTGTGTCTCATCGTCACCGTCCGCGCGCGCGCAAGCTGTTC 1223
Qy 323 ValLeuSerGlyValValMetValPheSerThrAlaPheGlyAlaTrpPheLysLeu 342
Db 1224 CTCAGGCGCGCGCAGATGATCGTGTGCGACGTATCGTGGCGAGCTCATCGCGCC 1283
Qy 343 ThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAlaGln 362
Db 1284 AAGTTCGGGACCGCGCACGGG----- 1307
Qy 363 ProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIle 382
Db 1308 -----GACATCGCAAGGGGTACGCGCGGTCTGTGTGTTTCATCTCGCGCTACGTC 1361

RESULT 6

US-09-679-686B-15
; Sequence 15, Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
; FILE REFERENCE: B3160 US NA
; CURRENT APPLICATION NUMBER: US/09/679,686B
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/081,131
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: PCI/US99/07561
; PRIOR FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1960
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (10)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (28)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (514)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (533)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (542)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (552)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (620)
; OTHER INFORMATION: n = a, c, g or t
US-09-679-686B-15

Alignment Scores:

Pred. No.: 9,02e-44 Length: 1960
 Score: 518.00 Matches: 148
 Percent Similarity: 47.5% Conservative: 95
 Best Local Similarity: 28.96% Mismatches: 198
 Query Match: 21.08% Indels: 70
 DB: 4 Gaps: 16

US-09-886-954A-1 (1-477) x US-09-679-686B-15 (1-1960)

```

QY      7  GluGluThrGlnProLeuLeuGlyPro-----ProGly 17
DB      141 GAGAAAAAATGCGCTGCGGTAGAGGTATTAGCAACGAGGGGCAAGGAGTACCTGGA 200

QY      18  GlySerAlaProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuGlyPro 37
DB      201 AGCCTCACTCCT-----TTGTGACGGTAACATGTATAGTTGCAGCCATGGGTGGG 251

QY      38  LeuSerPheGlyPheAlaLeuGlyTyrSerSerProAlaIle-----ProSerLeu 54
DB      252 TTAATCTTCGGTTACATATCGGAATTTTCAGGTGGGTGACATCCATGATCCGTTTCG 311

QY      55  GlnArgAlaAlaProAlaProArg-----63
DB      312 CTCAGATTTTCCCGTCGGTTCGGAAGAAGAAATTCGACAAAACGGTGAACCGATAC 371

QY      64  -----LeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAla 81
DB      372 TGTCATACGACAGTCAGACATGACGATGTCACGTCGCTGCTGTATCTCGCGCGTTG 431

QY      82  AlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlyLeuSerLeuLeu 101
DB      432 CTGTCGTGTTGGTTCCTCCACCGTCACAGTAGTTTCGGCGGAAATCTCCATGCTT 491

QY      102  LeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrp 121
DB      492 TTCGAGGCTTGGCTTTTCTCGCGGTGCTTATCAACGGNTTTCGCCANCAACGTTGG 551

QY      122  MetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAla 141
DB      552 NTGCTATCGTGGTGGATCTGCTCGGGTTCGGTATCCGGTTTGGCAATCAGTCTGTG 611

QY      142  ProValTyrIleSerGluIleAlaTrpProAlaValArgGlyLeuLeuGlySerCysVal 161
DB      612 CCACCTTANCTATCTGAATGCTCCATACAAATATAGAGAGCATTTGAACATGGCTTT 671

QY      162  GlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu--- 180
DB      672 CAGTTGTCCATCACTGTGTGTATCTGTGCGCAATGTGTGAACATATTTCTTTGCTAAA 731

QY      181  -----TrpArgTrp---LeuAlaValLeuGlyCysValProProSerLeuMet 195
DB      732 ATCAAGAGTGTGGGATGAGGTGATGTTGGAGGTGCTATGGTCCCTGCCCTTATA 791

QY      196  LeuLeuLeu---MetCysPheMetProGluThrProArgPheLeuLeuThrGlnHisArg 214
DB      792 ATCAGTAGGATCACTAGTCTTCCACACTCCCAATCCATGATGAAGGGGTGAT 851

QY      215  ArgGlnGluAlaMetAlaAlaLeuArgPheLeuTrpGly-----SerGluGlnGly 231
DB      852 CGCAGAGGCGCAAGGCTCAGCTTCAGAGAATTCGCGCATCGCAATGTTGATGAAGAG 911

QY      232  TrpGluAspProProIleGlyAlaGluGlnSerPheHisLeuAla-----246
DB      912 TTCATGACCTTGTGGCAGCAAGTGAATCTCTAGCCCAAGTGGAGCCCTTGGAGGAAC 971

QY      247  LeuLeuArgGlnProGlyIleTyrLysProPheIle---IleGlyValSerLeuMetAla 265
DB      972 TTGTTGCAAGAAAG-----TACAGACCCCACTCACCATGGCAGGTGATTCATTC 1025

QY      266  PheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIlePheGluGlu 285
DB      1026 TTCAGCAACTCACTGGAATCAATGTATCATGTATTTTACGCGCGGCTTTCAGTCTC 1085
  
```

```

QY      286  AlaLysPheLysAspSerSer-----LeuAlaSerValValValGlyValIleGlnVal 303
DB      1086 ATCGGGTTTAAGGATGATGCTCTCTAATGTGAGCTGTGATCACCGCGTGTCTTAATGTT 1145

QY      304  LeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuVal 323
DB      1146 GTCCCAACTTGTGTCTCAATTTATGTTGTGACAAGTGGGTAGGAGGCCCTTTTCCTT 1205

QY      324  LeuSerGlyValValMetValPheSer-----ThrSerAlaPheGlyAlaTyr 339
DB      1206 GAAGTGGAGTCCAAATGCTCATTTGCCAGGCTGTAGTTGAGCTGCAATTGGAGCAAG 1265

QY      340  PheLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProVal 359
DB      1266 TTTGGA---ACTGATGGGAACCCAGT-----1289

QY      360  SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCys 379
DB      1290 -----GATTTGCCAAAGTGGTATGCAATGTTGTGGTTCCTCTCATTTGC 1334

QY      380  LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGlu 399
DB      1335 ATTTATGTATCAGCATTTGCTGTCATGGGTCCCTAGGTGGTGGTGGTGGTGGTGGT 1394

QY      400  IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeu 419
DB      1395 ATCTTCCCTTGGAGATTCGTTTCAGCTGCTCAGAGTATCAATGTCGTTGCAACATGCTT 1454

QY      420  MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGly 439
DB      1455 TTCATTTCTTGTATGCAAGTCTTCTTGCAATGCTTTGCCACATGAAG---TTCGGC 1511

QY      440  AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
DB      1512 TTGTTCTCTCTTCTTGGCTTCTGCTGTGATCATGATCATCTTCTGCTACTTCTTCTT 1571

QY      460  ProGluThrLysGlyLysThrLeuGluGlnIle 470
DB      1572 CCCGAAACAAAGGCATTTCCAATTGAAGAAATG 1604
  
```

RESULT 7

US-09-643-597-135
 ; Sequence 135, Application US/09643597
 ; Patent No. 6426072
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Li, Samuel X.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: McNeill, Patricia D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121.455C11
 ; CURRENT APPLICATION NUMBER: US/09/643,597
 ; NUMBER OF SEQ ID NOS: 369
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 135
 ; LENGTH: 2856
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; US-09-643-597-135

Alignment Scores:
 Pred. No.: 3,76e-43 Length: 2856
 Score: 514.50 Matches: 145

```
Percent Similarity: 46.92%      Conservative: 91
Best Local Similarity: 28.82%    Mismatches: 168
Query Match: 20.94%             Indels: 99
DB: 4                            Gaps: 14

US-09-886-954A-1 (1-477) x US-09-643-597-135 (1-2856)

Qy 25 ArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeu 44
Db 210 GCGCTCATGCTGGCTGGAGAGCAGTGTGGCTCCCTGCGAGTTTGGCTACACACT 269
Qy 45 Gly----- 45
Db 270 GGAGTCATCAATGCCCCCGAGAGGTGATCGAGGAGTTCTACAACAGACATGGGTCCAC 329
Qy 46 ---TyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArgLeu 64
Db 330 CGCTATGGGGAGAGCATCTGCCACC----- 356
Qy 65 AspAspAlaAlaAsaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAlaGlyGly 84
Db 357 ---ACGCTCACACGCTCTGGTCCCTCTCAGTGGCATCTTT-----TCTGTTGGGGC 407
Qy 85 ValLeuGly-----GlyTrpLeuValAspArgAlaGlyArgLeuSerLeu 100
Db 408 ATGATGGCTCCTCTCTGTGGGCTTTTGTAAACGCTTTGGCCGGCGGAATCAATG 467
Qy 101 LeuLeuCysSer---ValProPheValAla-----GlyPheAlaValIleThr 115
Db 468 CTGATGATGAACCTGCTGGCTTCTGTGTCGGCGGTGCTCATGGGCTTCTCGAACTG--- 524
Qy 116 AlaAlaGlnAspValTrpMetLeuGlyArgLeuLeuThrGlyLeuAlaCysGly 135
Db 525 ---GGCAGTCTTTGAGATGCTGATCTGGGCCGCTTCTCATGGGCTTCTCGAACTG--- 581
Qy 136 ValAlaSerLeuValAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGly 155
Db 582 CTGACCACAGGCTCGTGCCCATGATGTGGGTGAAGTGTACCCACAGCCTTTCTGTGG 641
Qy 156 LeuLeuGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAla 175
Db 642 GGCCTGGGACCTCGACCAGCTGGGCATCGTCTGGGCATCTCTCGGCATCTCTGCCAGGTGTC 701
Qy 176 GlyTrpValLeuGlu-----TyrArgTrpLeuAlaValLeu 187
Db 702 GGC-----CTGACATCCATCATGGCAACAGGACCTGTGCCCCCTGCTGTGAGCATC 755
Qy 188 GlyCysValProProSerLeuMetLeuLeuMetCysPheMetProGluThrProArg 207
Db 756 ATCTTCATCCGGCCCTGTGTCAGTGCATCGTGTGCCCTTCTGCCCGGAGAGTCCCGC 815
Qy 208 PheLeuLeu---ThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuTrp 226
Db 816 TTCCTGCTCATCAACGCCAACGAGGAGAACCGGGCCAGAGGTGTGCTAAGAAGCTGGCC 875
Qy 227 GlySerGluGlnGlyTrpGluAspProProIleGlyAlaGluGlnSerPheHis----- 244
Db 876 GGGACAGCTGAGTGCATCCCATGACCTGCGAGGAGATGAAGAGAGATCGCGAGATGATG 935
Qy 245 -----LeuAlaLeuLeuArgGlnProGlyIleTyrIysProPhe 257
Db 936 CGGGAGAAAGAGTCAACCATCTCTGGAGCTGTTCGGCTCCCGCCCTACCGCCAGCCCATC 995
Qy 258 IleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPhe 277
Db 996 CTCATCGCTGTGGTGTGTCAGTGTCCCGAGCTGTCTGGCATCAACCTGTCTTCAT 1055
Qy 278 TyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValVal 297
Db 1056 TACTCCACAGCATCTTCGAGAGGGGGGGTGCACGAGCCTGTGTATGCCACCATGTC 1115
Qy 298 ValGlyValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGly 317
```

```
Db 1116 TCCGCTATCGTCAACACGGCCTTCACTGTCGTGCTGTTGTGTGGAGCAGCAGGC 1175
Qy 318 ArgArgLeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGly 337
Db 1176 CGCGGACCTGCGACCTCATTA----- 1196
Qy 338 AlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAla 357
Db 1197 -----GGCCTCGCTGCATGGCGGTTGTGCCATCATCTCATG 1232
Qy 358 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySer 377
Db 1233 ACCATCGCGCTAGCACTGCTGCAGCAGCTACCCCTGGATGCTCTATCTGAGCATGTGGCC 1292
Qy 378 MetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMet 397
Db 1293 ATCTTGGCTTGTGGCTTCTTTGAAGTGGTCTGCCCCCATCCCATGGTTCATCGTG 1352
Qy 398 SerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsn 417
Db 1353 GCTGAACCTCTTCAGCCAGGTCACGTCAGCTGCCATTGCGGTTCAGAGCTTCTCCAAC 1412
Qy 418 TrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro 437
Db 1413 TGGACCTCAAAATTTCAATTGTCGGCATGTCTCCAGTATGTGGAGCAACTGTGTGTC 1472
Qy 438 TyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu----- 453
Db 1473 TAC-----GTCTTCATCATCTTCACCTGTCTCTGTTCTGTTC 1511
Qy 454 -----PheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThr 471
Db 1512 TTCATCTTCACCTACTTCAAAGTTCCTGAGACTAAAGCCGACCTTCGATGAGATCGCT 1571
Qy 472 AlaHisPhe 474
Db 1572 TCCGGCTTC 1580

RESULT 8
US-09-480-884A-135
; Sequence 135, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-480-884A-135

Alignment Scores:
Pred. No.: 3,76e-43      Length: 2856
Score: 514.50           Matches: 145
Percent Similarity: 46.92%      Conservative: 91
Best Local Similarity: 28.82%    Mismatches: 168
Query Match: 20.94%           Indels: 99
DB: 4                    Gaps: 14

US-09-886-954A-1 (1-477) x US-09-480-884A-135 (1-2856)

Qy 25 ArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeu 44
```

Db 210 CCGCTCATGCTGGCTGTGGAGGAGCAGTGTGGCTCCCTCGAGTTTGGCTACAACACT 269
 QY 45 Gly----- 45
 Db 270 GGAGTCATCAATGCCCGCCAGAGGTGATCGAGAGTTCTACAACAGACATGGGTCCAC 329
 QY 46 ---TyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArgLeu 64
 Db 330 CGCTATGGGAGAGCATCTGCCACC----- 356
 QY 65 AspAspAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaGlyGly 84
 Db 357 ---ACGCTCACACCGCTCTGTGTCCTCTCAGTGGGCATCTT-----TCTGTGGGGGC 407
 QY 85 ValLeuGly-----GlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 Db 408 ATGATGGCTCCTTCTCTGTGGCCCTTTTCGTTAACCGCTTTGGCGCGGGAATTCAATG 467
 QY 101 LeuLeuCysSer---ValProPheValAla-----GlyPheAlaValIleThr 115
 Db 468 CTGATGATGAACCTGCTGGCCCTTCGTGCGCGGTGCTCATGGGCTTCTGNAACGTG--- 524
 QY 116 AlaAlaGlnAspValTrpMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGly 135
 Db 525 ---GGCAAGTCCCTTTGAGATGCTGATCTCTGGCGCGCTTCATCATCGGTGTGACTCGCGC 581
 QY 136 ValAlaSerLeuValAlaProValTrpIleSerGluIleAlaTyrProAlaValArgGly 155
 Db 582 CTGACCACAGGCTTCGTCCTCATGATGTGGGTGAAGTGTACCCACAGCCCTTCGTGGG 641
 QY 156 LeuLeuGlySerCysValGlnLeuMetValValGlyIleLeuLeuAlaTyrIleAla 175
 Db 642 GCCCTGGGCACCCCTGACACAGTGGCATCTGCTCGCATCCTCATCGCCAGGTGTTC 701
 QY 176 GlyTrpValLeuGlu-----TrpArgTrpLeuAlaValLeu 187
 Db 702 GGC-----CTGAGTCCATCATGGCAACAGACCTGTGGCCCTGCTGCTGAGCATC 755
 QY 188 GlyCysValProProSerLeuMetLeuLeuMetCysPheMetProGluThrProArg 207
 Db 756 ATCTTCATCCCGCCCTGTCAGTGCATCTGCTGTCCTTCTGCCCGAGAGTCCCGC 815
 QY 208 PheLeuLeu---ThrGlnHisArgArgGlnGlnAlaMetAlaAlaLeuArgPheLeuTrp 226
 Db 816 TTCCTGCTCATCAACCGCAACAGAGAACCGGCCAAGAGTGTCTAAAGAGCTGGC 875
 QY 227 GlySerGluGlnGlyTrpGluAspProPheGlyAlaGluGlnSerPheHis----- 244
 Db 876 GGGACAGCTGACGTGACCCATGACCTGCAGGAGATGAAGGAAGAGAGTCCGACAGATG 935
 QY 245 -----LeuAlaLeuLeuArgGlnProGlyIleTyrLysProPhe 257
 Db 936 CGGAGAGAAGAGTCAACATCTGGAGCTGTCCGCTCCCGCCGCTACCGCACGCCCATC 995
 QY 258 IleIleGlyValSerLeuMetAlaPheGlnLeuSerGlyValAsnAlaValMetPhe 277
 Db 996 CTATCGCTGGGCTGTCACCTGCCAGAGCTGTGTGCATCAACGCTGTCTTCTAT 1055
 QY 278 TyrAlaGluThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSerValVal 297
 Db 1056 TACTCCACGAGCATCTTCGAGAGGGGGGTGCAGCAGCCCTGTGTATGCCACCATTTGGC 1115
 QY 298 ValGlyValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGly 317
 Db 1116 TCCGGTATCGTCAACAGCGCTTCACGTCTGCTGTGCTGTGTGTGGTGGAGCGAGCAGC 1175
 QY 318 ArgArgLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGly 337
 Db 1176 CGCGGACCCCTGCACCTATA----- 1196
 QY 338 AlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAla 357
 Db 1197 -----GGCTCGCTGGCATGGGGTGTGGCCATCTCATG 1232

QY 358 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpPheValAlaValGlySer 377
 Db 1233 ACCATCGCGCTAGCACTGCTGGAGCAGCTACCCCTGGATGCTCTATCTGAGCATCGTGGCC 1292
 QY 378 MetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMet 397
 Db 1293 AFCTTTGGCTTGTGGCCTTCTTTGAAGTGGGTCTCTGGCCCATCCCATGGTTTCATCGTG 1352
 QY 398 SerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsn 417
 Db 1353 GCTGAACCTTTCAGCCAGGTCAGCTCCAGCTGCATGCGTTCAGGCTTCTCCAAAC 1412
 QY 418 TrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro 437
 Db 1413 TGGACCTCAAAATTTTCATTGTGGCATGTCTTCCAGTATGTGGAGCAACTGTGTGTCCTCC 1472
 QY 438 TyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu----- 453
 Db 1473 TAC-----GTCCTTCATCATCTTCACTGTCTCTCTGTTCTGTTC 1511
 QY 454 -----PheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThr 471
 Db 1512 TTCATCTTTCACCTACTTCAAGTTCCTGAGACTAAAGCCGACCTTCGATGAGATCGCT 1571
 QY 472 AlaHisPhe 474
 Db 1572 TCCGGCTTC 1580

RESULT 9

US-09-542-615A-135
 ; Sequence 135, Application US/09542615A
 ; Patent No. 6518256
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Ligu
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy A.
 ; APPLICANT: Fanger, Gary R.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
 ; FILE REFERENCE: 210121.455C8
 ; CURRENT APPLICATION NUMBER: US/09/542.615A
 ; CURRENT FILING DATE: 2000-04-14
 ; NUMBER OF SEQ ID NOS: 350
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 135
 ; LENGTH: 2856
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-542-615A-135

Alignment Scores:
 Pred. No.: 3,76e-43 Length: 2856
 Score: 514.50 Matches: 145
 Percent Similarity: 46.92% Conservative: 91
 Best Local Similarity: 28.83% Mismatches: 168
 Query Match: 20.94% Indels: 99
 DB: 4 Gaps: 14

US-09-886-954A-1 (1-477) x US-09-542-615A-135 (1-2856)

QY 25 ArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeu 44
 Db 210 CGCTCATGCTGGCTGGGAGGAGCAGTCTTGGCTCCCTGACGTTTGGCTACACACT 269
 QY 45 Gly----- 45
 Db 270 GGAGTCATCAATGCCCGCCAGAGGTGATCGAGAGTTCTACAACAGACATGGGTCCAC 329
 QY 46 ---TyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArgLeu 64

330 CGCTATGGGAGAGCATCTGCCACCC----- 356
QY 65 AspaepAlaAlaAAserTrpPheGlyAlaValValThrLeuGlyAlaAlaAAGlyGly 84
Db 357 ---ACGCTCACACAGCTCTGGTCCCTCTCAGTGGCCATCTTT-----TCTGTTGGGGC 407
QY 85 ValLeuGly-----GlyTrpLeuValAspArgAlaGlyValGlyLeuSerLeu 100
Db 408 ATGATTTGGCTCTCTCTGTGGCCCTTTTCGTTAAACCGCTTGGCCGGCGGAAATCAATG 467
QY 101 LeuLeuCysser---ValProPheValAla-----GlyPheAlaValIleThr 115
Db 468 CTGATGATGAACCTCTGGCTTCTGTGCGCGCGTCTCATGGGCTTCTCGAAACAGT--- 524
QY 116 AlaAlaGlnAspValTrpMetLeuLeuGlyGlyArgLeuThrGlyLeuAlaCysGly 135
Db 525 ---GGCAAGTCTTTGAAGATGCTGATCCTGGCCCGCTTCATCATCGGTGTACTGCGGC 581
QY 136 ValAlaSerLeuValAlaProValTrpIleSerGluIleAlaTyProAlaValArgGly 155
Db 582 CTGACCACAGGCTTCGTGCCATGATGTGGGTGAAGTGTACCCACAGCTTTCGTGGG 641
QY 156 LeuLeuGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAlaTyTrpLeuAla 175
Db 642 GCCCTGGGCACCTGCACAGCTGGGCATCGTGTGGCATCTCTCATGCCAGGTGTC 701
QY 176 GlyTrpValLeuGlu-----TrpArgTrpLeuAlaValLeu 187
Db 702 GGC-----CTGACTCCATCATGGGCAACAGGACCTGTGGCCCTCTGCTGAGCATC 755
QY 188 GlyCysValProSerLeuMetLeuLeuMetCysPheMetProGluTrpProArg 207
Db 756 ATCTTCATCCCGCCCTCTGTGAGTCATCGTGTCCCTTCTGCCCGGAGAGTCCCGC 815
QY 208 PheLeuLeu---ThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuTrp 226
Db 816 TTCCTGCTCATCAACCGCAACGAGGAGAACCGGGCAAGAGTGTGTTAAGAAGCTGGC 875
QY 227 GlySerGluGlnGlyTrpGluAspProIleGlyAlaGluGlnSerPheHis----- 244
Db 876 GGGACAGCTGACGTGACCCATGACCTGCGAGGAGATGAAGGAAGAGAGTGGCGAGATG 935
QY 245 -----LeuAlaLeuLeuArgGlnProGlyIleTyTrpLeuProPhe 257
Db 936 CGGGAGAAGAGGTACCATCTCTGGAGCTGTCCGCTCCCGCTACCGCCAGCCCATC 995
QY 258 IleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPhe 277
Db 996 CTCATCGCTGTGTGTCAGCTGTCCAGCAGCTGTCTGGCATCAACGCTCTCTCTAT 1055
QY 278 TyrAlaGluThrIlePheGluGluAlaTyPheLysAspSerSerLeuAlaSerValVal 297
Db 1056 TACTCCACGAGATCTTCGAGAAGGGGGGTGCGACGACCTGTGTATGCCACCATGGC 1115
QY 298 ValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGly 317
Db 1116 TCCGGTATGCTCAACACGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1175
QY 318 ArgArgLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGly 337
Db 1176 CGGCGGACCTGACCTCATA----- 1196
QY 338 AlaTyTrpPheIleLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAla 357
Db 1197 -----GGCTCGCTGGCATGGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1232
QY 358 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySer 377
Db 1233 ACCATCGCGCTAGCACTGTCTGGAGCAGCTACCCCTGGATGCTCTATCTGAGCATCGTGGC 1292
QY 378 MetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuMet 397
Db 1293 ATCTTTGGCTTTGTGGCTTCTTTGAAGTGGGTCTCTGGGCCCCATCCATGGTTCATCGT 1352

QY 398 SerGluIlePheProLeuHisValIysGlyValAlaThrGlyIleCysValLeuThrAsn 417
Db 1353 GCTGAACCTCTCAGCCAGGTCCACGTCACCTGCCATTCGCTTGGAGGCTTCTCOAAC 1412
QY 418 TrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro 437
Db 1413 TGGACCTCAAAATTTCAATTTGGGCATGTCTTCCAGTATGTGGAGCAACTGTGTGGTCCC 1472
QY 438 TyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu----- 453
Db 1473 TAC-----GTCTTCATCACTTCATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1511
QY 454 -----PheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThr 471
Db 1512 TTCATCTTCACCTACTTCAAGTCTCTGAGACTAAAGCCGACCTTCGATGAGATCGCT 1571
QY 472 AlaHisPhe 474
Db 1572 TCCGGCTTC 1580

RESULT 10
US-09-606-421B-135
; Sequence 135, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangor, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C9
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-606-421B-135

Alignment Scores:
Pred. No.: 3,76e-43 Length: 2856
Score: 514.50 Matches: 145
Percent Similarity: 46.92% Conservative: 91
Best Local Similarity: 28.83% Mismatches: 168
Query Match: 20.94% Indels: 99
DB: 4 Gaps: 14

US-09-886-954A-1 (1-477) x US-09-606-421B-135 (1-2856)

QY 25 ArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeu 44
Db 210 CGCCTCATGCTGGCTGTGGGAGGAGCGATGCTGGCTCCCTGCTGCTGCTGCTGCTGCT 269
QY 45 Gly----- 45
Db 270 GGAGTCATCAATGCCCCCAGAGGTGTCGAGGAGTCTACACACAGACATGGGTCAC 329
QY 46 ---TyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProAlaProArgLeu 64
Db 330 CGCTATGGGAGAGCATCTCTGCCACC----- 356
QY 65 AspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaGlyGly 84
Db 357 ---ACGCTCACACAGCTCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 407

QY 85 ValLeuGly-----GlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db : : : : :
408 ATGATGGTCCCTCTCTGTGGCCCTTTTCGTTAACCGCTTTGGCCGGCGGAATCAATG 467
QY 101 LeuLeuCysSer---ValProPheValAla-----GlyPheAlaValIleThr 115
Db : : : : :
468 CTGATGATGAACCTGTCGCTTCGTGTCGCCGTCTCATGGCTTCTCGAAACTG--- 524
QY 116 AlaAlaGlnAspValTrpMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGly 135
Db : : : : :
525 ---GGCAAGTCCTTTGAGATGCTGATCCTGGCCGCTTCATCATCGGTGTGTACTGCGGC 581
QY 136 ValAlaSerLeuValAlaProValTrpIleSerGluIleAlaTyrProAlaValArgGly 155
Db : : : : :
582 CTGACACAGGCTTCGTGCCCATGATGTGGTGAAGTGTACCCACAGCCCTTCGTGGG 641
QY 156 LeuLeuGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAla 175
Db : : : : :
642 GCCCTGGGCACCTGCACAGCTGGGCATGTCGTGCGCATCTCATGCCCGAGGTTC 701
QY 176 GlyTrpValLeuGlu-----TrpArgTrpLeuAlaValLeu 187
Db : : : : :
702 GGC-----CTGGACTCCATCATGCGCAACAGGACCTGTGGCCCTCTGTGAGCATC 755
QY 188 GlyCysValProProSerLeuMetLeuLeuMetCysPheMetProGluThrProArg 207
Db : : : : :
756 ATCTTCATCCGCCCTCTGTCAGTGCATGTCGTGCTGCTCCCGAGAGTCCCCGC 815
QY 208 PheLeuLeu---ThrGlnHisArgArgGlnGluAlaMetAlaLeuArgPheLeuTrp 226
Db : : : : :
816 TTCCTGTCTCATCAACCGCAACAGAGAACCGGGCAAGAGTGTCTAAAGAGCTGGC 875
QY 227 GlySerGluGlnGlyTrpGluAspProProlleGlyAlaGluGlnSerPheHis----- 244
Db : : : : :
876 GGGACAGTACGTCGACCCATGACCTGCAGGAGATGAAGAGAGAGTGGCAGATGATG 935
QY 245 -----LeuAlaLeuLeuArgGlnProGlyIleTyrLysProPhe 257
Db : : : : :
936 CGGGAGAAGAGTCAACCATCTGGAGTGTTCGCTCCCGCTCCCGCTCCCGAGAGTCC 995
QY 258 IleIleGlyValSerLeuMetAlaPheGlnLeuSerGlyValAsnAlaValMetPhe 277
Db : : : : :
996 CTCATCGT 1055
QY 278 TyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerVal 297
Db : : : : :
1056 TACTCCAGCAGCATCTTCGAGAAGCGGGGTGCAGCAGCTGTGTGTGTGTGTGTGTGT 1115
QY 298 ValGlyValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGly 317
Db : : : : :
1116 TCCGGTATCGTCAACACCGCTTCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1175
QY 318 ArgArgLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGly 337
Db : : : : :
1176 CGGGGAGCCCTGCACCTCATA----- 1196
QY 338 AlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAla 357
Db : : : : :
1197 -----GGCTCGCTGGCATGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1232
QY 358 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySer 377
Db : : : : :
1233 ACCATCGGCTAGCAGTCTGGACAGTACCTCGGATGCTCTATCATGAGATCGTGGCC 1292
QY 378 MetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMet 397
Db : : : : :
1293 ATCTTTGGCTTTGGGCTTCTTTGAAGTGGTCTCTGGCCCACTCCCATGTTTCATGTG 1352
QY 398 SerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsn 417
Db : : : : :
1353 GCTGAACCTCTTCACCGAGGTTCCACGTCCTCCATTCGCGTTGCGGCTTCTCCAC 1412

QY 418 TrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro 437
Db : : : : :
1413 TGGACCTCAAAATTTTCATTTGGGCGATGCTTCCAGTATGTGAGCAACTGTGTGTGCC 1472
QY 438 TyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu----- 453
Db : : : : :
1473 TAC-----GTCTTCATCACTTCTACTGTCTCTCTGTGTCTGTCTGTCTGTCT 1511
QY 454 -----PheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThr 471
Db : : : : :
1512 TTCATCTTCACTTCAAGATTCTCAGACTAAAGCCGCGACCTTCGATGAGATCGCT 1571
QY 472 AlaHisPhe 474
Db : : : : :
1572 TCCGGCTTC 1580
RESULT 11
US-09-221-107-135
; Sequence 135, Application US/09221107
; Patent No. 6660838
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121, 455C2
; CURRENT APPLICATION NUMBER: US/09/221,107
; CURRENT FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Human
US-09-221-107-135
Alignment Scores:
Pred. No.: 3,76e-43 Length: 2856
Score: 514.50 Matches: 145
Percent Similarity: 46.92% Conservative: 91
Best Local Similarity: 28.83% Mismatches: 168
Query Match: 20.94% Indels: 99
DB: 4 Gaps: 14
US-09-886-954A-1 (1-477) x US-09-221-107-135 (1-2856)
QY 25 ArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeu 44
Db : : : : :
210 CGCTCATGT 269
QY 45 Gly----- 45
Db : : : : :
270 GGAGTCATCAATGCCCCCAGAGGTGATCGAGAGTTCTAACACAGACATGGGTCCAC 329
QY 46 ---TyrSerSerProAlaIleProSerLeuGlnArgAlaProProAlaProArgLeu 64
Db : : : : :
330 CGCTATCGGGAGAGCATCTGCCCCACC----- 356
QY 65 AspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaGlyGly 84
Db : : : : :
357 ---ACGCTCACCAGCTCTGCTCCCTCTCATGGCCATCTT-----TCTGTGGGGGC 407
QY 85 ValLeuGly-----GlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db : : : : :
408 ATGATGGTCCCTCTCTGTGGGCTTTTCGTTAACCGCTTTGGCCGGCGGAATCAATG 467
QY 101 LeuLeuCysSer---ValProPheValAla-----GlyPheAlaValIleThr 115
Db : : : : :
468 CTGATGATGAACCTGTCGCTTCGTGTCGCCGTCTCATGGCTTCTCGAAACTG--- 524
QY 116 AlaAlaGlnAspValTrpMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGly 135
Db : : : : :
525 ---GGCAAGTCCCTTTGAGATGCTGATCCTGGCCGCTTCATCATCGGTGTGTACTGCGGC 581
QY 136 ValAlaSerLeuValAlaProValTrpIleSerGluIleAlaTyrProAlaValArgGly 155

```
Db 582 CTGACCAGAGCTTCGTCCCATGTATGTGGTGAAGTGTACCCACAGCCTTTCTGTGG 641
Qy 156 LeuLeuGlySerCysValGlnLeuMetValValValGlyLeuLeuAlaTyrLeuAla 175
Db 642 GCCTGGGACCTGCACACAGCTGGGCATCGTCGTGGCATCTCTCATGCCAGGTGTC 701
Qy 176 GlyTrpValLeuGlu-----TTPArgTrpLeuAlaValLeu 187
Db 702 GGC-----CTGGACTCCATATGGGCAACAGGACCTGTGGCCCTCTGCTGAGCATC 755
Qy 188 GlyCysValProSerLeuMetLeuLeuMetCysPheMetProGluTrpProArg 207
Db 756 ATCTCATCCGGCCCTGCTGCAGTCATCGTGTCCCTTCTGCCCGGAGAGTCCCGC 815
Qy 208 PheLeuLeu---ThrGlnHisArgArgGlnGluAlaMetAlaLeuArgPheLeuTrp 226
Db 816 TTCCTGCTCATCAACCGCAACGAGGAGAACCGGGCCAAAGAGTGTGCTAAAGAGCTGGC 875
Qy 227 GlySerGluGlnGlyTrpGluAspProProGlyAlaGlnGlnSerPheHis----- 244
Db 876 GGGACAGCTGAGCTGACCCATGACCTGCGAGGAGATGAAGGAAGAGAGTCCGCAGATGATG 935
Qy 245 -----LeuAlaLeuLeuArgGlnProGlyIleTyrLysProPhe 257
Db 936 CGGGAGAAGAGTCCACCATCTCGGAGCTGTTCGGCTCCCGCTACCGCCAGCCCATC 995
Qy 258 IleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPhe 277
Db 996 CTCATCGCTGTGTGTGTCAGCTGTCCACAGCTGTCTGGCATCAACGCTCTCTCTAT 1055
Qy 278 TyrAlaGluThrIlePheGluGlnAlaLysPheLysAspSerSerLeuAlaSerValVal 297
Db 1056 TACTCCAGGACATCTTCAGAAAGCGGGGTGCGAGCCCTGTGTATGCCACCATGGC 1115
Qy 298 ValGlyValIleGlnValLeuPheThrAlaValAlaLeuLeuMetAspArgAlaGly 317
Db 1116 TCCGGTATCGTCAACAGCGCTTCACTGCGTGTGCTGTGTGTGGTGGAGCGAGCGC 1175
Qy 318 ArgArgLeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGly 337
Db 1176 CGGCGGACCTGCACCTCAT----- 1196
Qy 338 AlaTyrPheLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAla 357
Db 1197 -----GCCCTCGCTGGCATGGCGGTGTGCCATACTCATG 1232
Qy 358 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpPheAlaValGlySer 377
Db 1233 ACCATCGCGCTAGCACTGTGGAGCAGCTACCTGCTATCTGATCATCGTGGCC 1292
Qy 378 MetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMet 397
Db 1293 ATCTTGGCTTTGTGGCCCTCTTTGAAGTGGGCTCTGGCCCCCATCCATGGTTCATCGT 1352
Qy 398 SerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsn 417
Db 1353 GCTGAACCTTTAGCCAGGTCACGTCCAGCTGCCATTGCCGTTCAGGCTTCTCCAC 1412
Qy 418 TrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro 437
Db 1413 TGGACCTCAAAATTTATTTGGGCATGTGCTTCCAGTATGTGGAGCAACTGTGTGGTCC 1472
Qy 438 TyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu----- 453
Db 1473 TAC-----GTCTTCATCATCTTCACTGTCTGTTCTGTTCTGTTCT 1511
Qy 454 -----PheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThr 471
Db 1512 TTCATCTTCACTTCTCAAGTTCCTGAGACTTAAGCGCGGACCTTCGATGAGATCGCT 1571
Qy 472 AlaHisPhe 474
::: |||
```

```
Db 1572 TCCGGCTTC 1580
RESULT 12
US-09-489-039A-4560
; Sequence 4560, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4560
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4560

Alignment Scores:
Pred. No.: 7,32e-43 Length: 1506
Score: 507.50 Matches: 149
Percent Similarity: 46.67% Conservative: 75
Best Local Similarity: 31.04% Mismatches: 188
Query Match: 20.66% Indels: 69
Db: 4 Gaps: 10

US-09-886-954A-1 (1-477) x US-09-489-039A-4560 (1-1506)
Qy 26 ValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPhe----- 42
Db 106 GTCTGAGCAATTGCTGTCGCCGCCCTCGCGTGTGTACTCTTTGGCTATGACTGGGTG 165
Qy 43 AlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProAlaPro 62
Db 166 GTGATTGGCGGCTAAGCCATTTTATGAAGCTGTTTCAATTACGACCCG----- 219
Qy 63 ArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAla 82
Db 220 -----GCCAGTCCGGCTGGCGATGAGCTACGCGCTGTGGGCTGTATTTC 267
Qy 83 GlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeu 102
Db 268 GCGCATTAATTTCCGGATGGTCCGACAACTGGGGCGCAGCTGCCATTATTCCT 327
Qy 103 CysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
Db 328 TCCGCGCTGCTCTTCAGCGCTGCGGCTGGGGAGCGGCGTCCGAGTCATTTCGATATG 387
Qy 123 LeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaPro 142
Db 388 TTGTGGTTTACCGATTGTGGCGGCGFAGGGATTGGTCTGGCTTCCCTCTCAGCCG 447
Qy 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
Db 448 CTTTACATTGCCAGTCAAGCCGGCAGAGAAAGAGGACGTTTGTCCGCGTCAATCAG 507
Qy 163 LeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu----- 180
Db 508 CTCACCATCGTATTGGCGTGTGGCCGCTCAGTTAATCAATCTGATGATCTGTAACCG 567
Qy 181 -----TTP 181
Db 568 GTGGAGCCGGGCGGCGAGCAGATGATTGTGGACAGCTGGAATGGGAGATGGGCTGG 627
Qy 182 ArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPhe 201
Db 628 CGCTGGATGTTCGGTCCGGAACCTGGTGTCCGCACTGGCGTTTCTGCTCTGATGTTT 687
Qy 202 MetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaAla 221
::: |||
```

```

Db 688 GTCCCGAGTCGCGCGTGGCTCATGAAGCGCGGTAAACCGAGCGCGCGCGGCTGG 747
QY 222 LeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspPro----- 235
Db 748 CTGGAAACGATGGTCTGCCGACTATGCCGA-CAGGATCCTCGTGAATCGCGCATAC 806
QY 236 ProIleGlyAlaGluGln-SerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIy 255
Db 807 CTGGAAAGGATAACATAAAGTCTCCTACGGCGGCTGTGGCTCCCGAGTGAACCC 866
QY 255 sProhellelleGlyValSerMetAlaPheGlnGlnLeuSerGlyValAsnAlaVa 275
Db 867 GATTGTGATCATGGCATGGTCTGCCATATTCCAGCAGTGGTGTGGGATTAAACGTAT 926
QY 275 lMetPheTyrAlaGluThrIlePheGluGluAlaIysPheLys---AspSerSerLeuAl 294
Db 927 CTTTAACTACGCGCGAGAGATTTTTCCTCGCGCGGGTTCGATATTAACAGCAGCTGAA 986
QY 294 aSerValVal---ValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMe 313
Db 987 ATCGATCGTCGACGCGCGTGTAACTGTCCTTACCATTGCGGCGCTGCCGCTGGT 1046
QY 313 tAspArgAlaGlyArgArgLeuLeuValLeu-----SerGlyValValMetValPh 331
Db 1047 GGATAAAATCGTCGCGCGTAAATTAATGCTGCTTGGCGCTTCGGGATTGACGCTGATCTA 1106
QY 331 eSerThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyGlyProGlyAsnSerSe 351
Db 1107 TGTGCTGATGCC---GGCGCTTACGCCATGGGCATTTATGGGGTGGCGG----- 1152
QY 351 rHisValAlaIleSerAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAl 371
Db 1153 -----GTACTGTTGCT 1163
QY 371 aTrpLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyPr 391
Db 1164 GGTGTCGCGCGG-----ATTGCTATTATGGTTGACCTGGCGCC 1205
QY 391 oIleProTrpLeuLeuMetSerGluIlePheProLeuHisValLysGlyValAlaThrGl 411
Db 1206 GGTGACCTGGGTGTCTGGCGGAGATTTTCCCAACCGCGTACGTGACTCGCGATGTC 1265
QY 411 yIleCysValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerSerLe 431
Db 1266 TTAGGTACCTGGCGCTGTGGATCGCTGTTCTCTGTAACTATATCTTCCCGCTGCT 1325
QY 431 uMetGluValLeuArgProTyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSe 451
Db 1326 TAATGCCGCTCTGGCGCGCGGAGAACTTCTGCTGTATGGTGTTATTTGGCGTGGCG 1385
QY 451 rValLeuPheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIle 470
Db 1386 CTACCTCTATATCTTGGCAACGTCCTCCGGAACGAAGGATATCACCTTCGAAGCGCTG 1443

```

RESULT 13

```

US-09-291-922-19
; Sequence 19, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tiney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19

```

```

; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Zea mays
US-09-291-922-19

```

Alignment Scores:

```

Pred. No.: 1,358-42 Length: 1914
Score: 506.50 Matches: 137
Percent Similarity: 46.25% Conservative: 91
Best Local Similarity: 27.79% Mismatches: 182
Query Match: 20.61% Indels: 83
DB: 4 Gaps: 9

```

US-09-886-954A-1 (1-477) x US-09-291-922-19 (1-1914)

```

QY 28 LeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGlyTyrSer 47
Db 122 ATATGTGCCATCTCGCTCCATGGCTCTGTCACTCTTGGCTATGACATTTGGGTGATG 181
QY 48 SerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArgLeuAspAla 67
Db 182 AGTGAGCGGCCATGTACATCAAGAAAGACCTG-----AATATACACGACGCTG 229
QY 68 AlaAlaSerTrpPheGlyAlaValThrLeuGlyAlaAlaAlaGlyGlyValLeuGly 87
Db 230 CAGCTGGAGATCTCGATCGGATCTCTAGTCTCTACTCGCTGTTCCGATCTCTCGCTGC 289
QY 88 GlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuLeuCysSerValProPhe 107
Db 290 GCGCGGACGTCGACAGGATCGGCGCGCTTGACCGTGGTTCGCGCTGTCTATCTTC 349
QY 108 ValIaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGlyArg 127
Db 350 TTCGTGGGCTCGTGTCTCATGGGTTTCGCCCTCAACTACGCGCATCTCATGCGGCGCG 409
QY 128 LeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaProValTyrIleSerGlu 147
Db 410 TTCGTGGCGGAGTGGGTGGCTACGGGGGCGATGATCGCGCGCTGTACACGCGCGAG 469
QY 148 IleAlaTrpProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetValValVal 167
Db 470 ATCTCGCTCGCGGCTCCCGTGGCTTCTGACACCTTCCCGGAGGTGTTCAACAATC 529
QY 168 GlyIleLeuLeuAlaTyrLeuAlaGlyTrpVal-----LeuGlu 180
Db 530 GGCACTCTCTGTGGCTACTGTCCAACCTTCGCTTCGCGCGCTCCCGCTCCACTCGCGC 589
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 590 TGGCGCTCATGCTCGCATTTGGCGAGTTCCGTCGCGCTGCTCGCGCTCTCGTGTTC 649
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
Db 650 TGCATGCCGAGTGCCTCGTGGTGTCTTTGAAGGGCGCTTCGCGGACGCCAGGCT 709
QY 221 AlaLeuArg----- 223
Db 710 GTGCTAGAGAAGACCTCTGCCACGAGAGGCGCGCGAGCGCTGGCGCATCAAG 769
QY 224 -----PheLeuTrpGlySerGlu 229
Db 770 GCGCGCGCGGGATTTCCGAAGGGCTCGAGGGGACGTAGTCACTACCGCGCAAGGAG 829
QY 230 GlnGly-----TrpGluAspProIleGlyAlaGluGlnSerPhe 243
Db 830 CAAGCGCGCGGTGAGTTGCAGGTCTGAAG----- 859
QY 244 HisLeuAlaLeuLeuArgGlnProGlyIleTyrIysProPheIleIleGlyValSerLeu 263
Db 860 AAGCTCATCTGTCCCGACCCCGGCTGTCCGAGCATACTGCTCTCGCGCGTGGGTCTC 919
QY 264 MetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIlePhe 283

```

```
Db 920 CACTTCTCCAGCAGGCTTCTGGCAGCGACTCCGTCGTCCAGTACAGCGCCGCGCTGTC 979
Qy 284 GluGluAlaLysPheGlyAspSerSerLeu-----AlaSerValValValGlyVal 300
Db 980 AAGAGCGGGGATCACCAGCAGCAACAAGCTCTCTGGGCGTCACCTGCGCGTGGCGGTG 1039
Qy 301 IleGlnValLeuPheThrAlaValAlaLysLeuMetAspArgAlaGlyArgLeu 320
Db 1040 ACCAAGAGCTTCTTCATCTGCTGGCCAGGTCCTGCTGGAGCCGCGGGCGTGGCGCT 1099
Qy 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaYrPhe 340
Db 1100 CTGCTGCTCATCAGCAGCGGGGATGATGCTCGCTCATCTGCTCGGTCG-----1153
Qy 341 LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaLysSerAlaProValSer 360
Db 1154 -----GGGCTCACCGTCGCG 1168
Qy 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrp-----LeuAlaValGlySer 377
Db 1169 GGGCATCACCGGACACCAAGTC-----GGTGGCGCGTCCCTGTGTCATCGGTCA 1222
Qy 378 MetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMet 397
Db 1223 ACCCTGTCATACGCTTCTCTCCATCGGCTCGGGCCATCAGCGGCGGTACACC 1282
Qy 398 SerGluLeuPheProLeuHisValLysGlyValAlaThrGlyLeCysValLeuThrAsn 417
Db 1283 TCGGAATATTCCCGCTGAGTGGCGCGCTGGGCTTGGCGGTGGGTGGCGGCAAC 1342
Qy 418 TrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro 437
Db 1343 CGGTCACCGCGCGCTCATCTCCATGACCTTCTCTGCTCCCTCCAGGCCATCACCATC 1402
Qy 438 TyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPhe 457
Db 1403 GCGCGCAGCTTCTTCTACTCTACGCGCATCGCGCGGTGCTTGGGTTTCTTCTTCTCAG 1462
Qy 458 CysValProGluThrLysGlyLysThrLeuGluGlnIle 470
Db 1463 TGCCTCCCGGACACGCGCGGACGCTGGAGGAGATG 1501
```

RESULT 14

```
US-09-591-025-8
; Sequence 8, Application US/09591025
; Patent No. 6303373
; GENERAL INFORMATION:
; APPLICANT: Bogan, Jonathan S.
; APPLICANT: Lodish, Harvey F.
; TITLE OF INVENTION: Method of Measuring Plasma Membrane
; TITLE OF INVENTION: Targeting of GLUT4
; FILE REFERENCE: 0399.1210-004
; CURRENT APPLICATION NUMBER: US/09/591,025
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/154,078
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/138,237
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 2592
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified GLUT4 containing myc tag sequences
US-09-591-025-8
```

Alignment Scores:

Pred. No.:	1-31e-41	Length:	2592
Score:	499.00	Matches:	155
Percent Similarity:	48.07%	Conservative:	82
Best Local Similarity:	31.44%	Mismatches:	173

```
Query Match: 20.31% Indels: 84
DB: 4 Gaps: 15
US-09-886-954A-1 (1-477) x US-09-591-025-8 (1-2592)
Qy 9 ThrGlnProLeuLeuGlyProProGlyGlySerAlaProArgGlyArgValPheLeu 28
Db 455 ACTCAAGACCTCTTG-----AGCCTGCAGAGGAGCAAAAGCTTATTCTG 499
Qy 29 AlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGlyTyrSerSer 48
Db 500 AAGAGGACTTGC-----TTAAGGAC 520
Qy 49 ProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArgLeuAspAlaAla 68
Db 521 CAGCT---CAATCCCTCCAGCACCTCACCACCTCTGG-GCCCTCTCCGTGGCCATC 576
Qy 69 AlaSerTrpPheGlyAlaValThrLeuGlyAlaAlaAlaGlyGlyValLeuGlyGly 88
Db 577 TTTTCGTCGGCGGCATGATT-----TCTCTCTCTCATTTGGTATCATCTCTCAG 627
Qy 89 TrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuLeuCysSerValProPheVal 108
Db 628 TGGCTT-----GGAAGGAAAGGGCCATGCTGGTCAACAATGTCTCTGCGCGTG 675
Qy 109 AlaGly-----PheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGly 125
Db 676 CTGGGGGCGAGCTCATGGGCTGGCCAAAGCTGTCTCTCTATGAATGTCATCTT 735
Qy 126 GlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaProValTyrIle 145
Db 736 GGCAGATTCTCATTTGGCGCTACTCAGGGCTGACATCAGGGCTGTGTCCTGACGTG 795
Qy 146 SerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetVal 165
Db 796 GGGGAGATTGCTCCCACTCACTGCGGGGCGCCCTGGGAGCGCTCAACAACCTGCCATT 855
Qy 166 ValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu-----180
Db 856 GTTATCGGCATCTCATCGCCAGGTGTGGGC-----TTGAGTCCCTCCTGGGCACT 909
Qy 181 -----TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeu 197
Db 910 GCGAGCTGTGGCCACTGCTCTCTGGGCTCAAGTGTACTGCTCCCTCCAGCTGGTGC 969
Qy 198 LeuMetCysPheMetProGluThrProArgPheLeu---LeuThrGlnHisArgGln 216
Db 970 CTGCTGCCCTTCTGTCCGAGAGCCCGCTACCTCTTACATCATCCAGAACTTCGAGGG 1029
Qy 217 GluAlaMetAlaAlaLeuArgPheLeu-----TrpGlySerGluGlnGly-----231
Db 1030 CCTGCCAGAAAGAGTCTGAAGCGCTGACAGGCTGGCGGATGTTCTGGAGTGTGGCT 1089
Qy 232 ---TrpGluAspProIleGlyAlaGluGlnSerPheHisLeuAlaLeuArgGln 250
Db 1090 GAGCTCAAGGATGAGAAGCGGAGCTGGAGCTGAGCGGCCACTGTCCCTGTCTCGCTC 1149
Qy 251 ProGlyIle-----TyrLysProPheIleIleGlyValSerLeuMetAlaPheGln 267
Db 1150 CTGGGCGAGCGGTACCCAGCGGAGCCCTGTATCTTGGTCTGCTGCTGAGCTGAGCCAG 1209
Qy 268 GlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIlePheGluAlaLys 287
Db 1210 CAGCTCTTGGCATCAATGCTGTTTCTATTATTTCGACCAGCATCTTCGAGACAGCAGG 1269
Qy 288 PheLysAspSerSerLeuAlaSerValValValGlyValIleGlnValLeuPheThrAla 307
Db 1270 GTAGGCCAGCCTGCTATGCCACCATGAGAGCTGGTGTGGTCAACACACTTTCACCTTG 1329
Qy 308 ValAlaAlaLeuIleMetAspArgAlaGlyArgLeuLeuLeuValLeuSerGlyVal 327
Db 1330 GTCTGGTGTGTTGTTGGTGGAGCGGGCGGCGCGGACCGCTCCATCTCTCTG-----1380
```



```
Db 1173 -----TGT 1175
QY 354 AlalleSerAlaProValSerAlaGlnProValAsp-AlaSerValGlyLeuAlaTrp-L 373
Db 1176 GCATCTTCATGTCAGTGGACTTGCTGCTGGAATAAGTTCTCTGGACTGAGTTACTG 1235
QY 373 euAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleP 393
Db 1236 TGAGCATATAGCCATCTCTCTTTGTCAGCTCTTTCAAAATTGGGCCAGGCCGATCC 1295
QY 393 roTrpLeuLeuMetSerGluIlePheProLeuHisValIysGlyValAlaThrGlyIleC 413
Db 1296 CCTGTTTCATGCTGCTGAGTTTTCAGTCAAGGACCCAGTCTGCTGCTTTAGCAATAG 1355
QY 413 ysValLeuThrAsnTrpLeuMetAlaPheLeuValThrIysGluPheSerSerLeuMetG 433
Db 1356 CTGCATTGAGCAATTGGACCTGCAATTTCATTGAGCTCTGTGTTTCCAGTACATTGGCG 1415
QY 433 luValLeuArgProTrpGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValL 453
Db 1416 ACTTCTGTGGACCTTATGTTTTCCTCTTTGCTGGAGTGCTCCTGGCCTTTACC--C 1472
QY 453 euPheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThrAlaH 473
Db 1473 TGTTTACATTTTAAAGTTCAGAAACCAAGGAAAGTCTTTTGAGGAAATTTGCTGCAG 1532
QY 473 isPheGlu 475
Db 1533 AATCCAA 1540
```

Search completed: September 28, 2004, 04:18:30
Job time : 152 secs